GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 10:50:49 ; Search time 22 Seconds
(without alignments)
81.435 Million cell updates/sec

Title: US-09-846-328B-1_COPY_2_25
Perfect score: 123
Sequence: 1 DAHKSEVAHRFKDLGEENFKALVL 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	123	100.0	28	4	US-09-846-329A-1	Sequence 1, Appli
2	123	100.0	585	1	US-08-153-799-14	Sequence 14, Appl
3	123	100.0	585	1	US-08-448-196A-3	Sequence 3, Appli
4	123	100.0	585	2	US-08-984-176-1	Sequence 1, Appli
5	123	100.0	585	2	US-08-702-572-2	Sequence 2, Appli
6	123	100.0	585	3	US-08-769-746-2	Sequence 2, Appli
7	123	100.0	609	1	US-08-222-619-3	Sequence 3, Appli
8	123	100.0	609	1	US-08-433-037-4	Sequence 4, Appli
9	123	100.0	609	4	US-08-897-956A-2	Sequence 2, Appli
10	123	100.0	609	4	US-09-976-594-977	Sequence 977, App
11	123	100.0	609	4	US-09-919-039-370	Sequence 370, App
12	123	100.0	609	5	PCT-US95-04075-3	Sequence 3, Appli
13	123	100.0	610	2	US-08-797-689-2	Sequence 2, Appli
14	123	100.0	610	4	US-09-984-186-2	Sequence 2, Appli
15	123	100.0	622	4	US-09-949-016-11170	Sequence 11170, A
16	123	100.0	783	1	US-08-256-938-2	Sequence 2, Appli
17	123	100.0	787	1	US-08-256-938-4	Sequence 4, Appli
18	123	100.0	787	2	US-08-797-689-16	Sequence 16, Appl
19	123	100.0	787	4	US-09-984-186-16	Sequence 16, Appl
20	123	100.0	978	4	US-08-897-956A-3	Sequence 3, Appli
21	109	88.6	582	1	US-08-134-638-1	Sequence 1, Appli
22	109	88.6	583	1	US-08-448-196A-4	Sequence 4, Appli
23	106	86.2	584	1	US-08-448-196A-7	Sequence 7, Appli
24	105	85.4	583	1	US-08-448-196A-6	Sequence 6, Appli
25	100	81.3	583	1	US-08-448-196A-5	Sequence 5, Appli
26	74	60.2	15	4	US-09-845-764A-1	Sequence 1, Appli
27	70	56.9	17	1	US-08-470-187-1	Sequence 1, Appli
28	70	56.9	17	1	US-08-318-905-1	Sequence 1, Appli
29	70	56.9	17	1	US-08-483-232-1	Sequence 1, Appli
30	70	56.9	17	1	US-08-483-140-1	Sequence 1, Appli
31	70	56.9	17	2	US-08-485-938A-1	Sequence 1, Appli
32	70	56.9	17	2	US-08-910-041-1	Sequence 1, Appli
33	70	56.9	17	3	US-09-328-474-1	Sequence 1, Appli
34	70	56.9	17	3	US-09-100-546-1	Sequence 1, Appli
35	70	56.9	17	3	US-09-010-715-1	Sequence 1, Appli
36	70	56.9	17	3	US-09-577-758-1	Sequence 1, Appli
37	68	55.3	13	2	US-08-803-364-7	Sequence 7, Appli
38	68	55.3	13	2	US-09-024-198-13	Sequence 13, Appl
39	68	55.3	13	2	US-09-186-409-13	Sequence 13, Appl
40	63	51.2	16	2	US-08-803-364-1	Sequence 1, Appli
41	63	51.2	16	2	US-09-024-198-10	Sequence 10, Appl
42	63	51.2	16	2	US-09-186-409-10	Sequence 10, Appl
43	60	48.8	13	2	US-08-803-364-6	Sequence 6, Appli
44	60	48.8	13	2	US-09-024-198-12	Sequence 12, Appl
45	60	48.8	13	2	US-09-186-409-12	Sequence 12, Appl

ALIGNMENTS

```
RESULT 1
US-09-846-329A-1
; Sequence 1, Application US/09846329A
; Patent No. 6620786
; GENERAL INFORMATION:
;  APPLICANT: Jackowski, George
;  TITLE OF INVENTION: Biopolymer Marker Indicative of Disease State Having A Molecular
;  TITLE OF INVENTION:  of 2937 Daltons
;  FILE REFERENCE: 2132.052
;  CURRENT APPLICATION NUMBER: US/09/846,329A
;  CURRENT FILING DATE:  2001-04-30
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 28
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-846-329A-1

  Query Match           100.0%;  Score 123;  DB 4;  Length 28;
  Best Local Similarity  100.0%;  Pred. No. 5.6e-13;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db          2 DAHKSEVAHRFKDLGEENFKALVL 25


RESULT 2
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
;  GENERAL INFORMATION:
;    APPLICANT:  Ballance, David J
;    APPLICANT:  Goodey, Andrew R
;    TITLE OF INVENTION:  Polypeptides
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  R Hain Swope, BOC Health Care Inc
;      STREET:  100 Mountain Avenue
;      CITY:  Murray Hill
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07974
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/153,799
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847975
;      FILING DATE:  06-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 8909916.2
;      FILING DATE:  29-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB90/00650
;      FILING DATE:  26-APR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/775952
;      FILING DATE:  29-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Swope, R Hain
;      REGISTRATION NUMBER:  24864
;      REFERENCE/DOCKET NUMBER:  92H832
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 665 2400
;      TELEFAX:  (908) 771 6159
;      TELEX:  219484
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  369..419
;      OTHER INFORMATION:  /note= "Alternative C-termini of
;      OTHER INFORMATION:  HSA(1-n)"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..585
;      OTHER INFORMATION:  /note= "Amino acid sequence of
;      OTHER INFORMATION:  natural HSA"
US-08-153-799-14

  Query Match             100.0%;  Score 123;  DB 1;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 3
US-08-448-196A-3
; Sequence 3, Application US/08448196A
; Patent No. 5780594
;  GENERAL INFORMATION:
;    APPLICANT:  CARTER, DANIEL C.
;    TITLE OF INVENTION:  BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
;    TITLE OF INVENTION:  CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
;    TITLE OF INVENTION:  RELATED PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NASA
;      STREET:  MARSHALL SPACE FLIGHT CENTER
;      CITY:  HUNTSVILLE
;      STATE:  ALABAMA
;      COUNTRY:  USA
;      ZIP:  35812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,196A
;      FILING DATE:  23-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROAD JR., ROBERT L.
;      REGISTRATION NUMBER:  18,757
;      REFERENCE/DOCKET NUMBER:  XX/MFS-28402-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  205-544-0021
;      TELEFAX:  205-544-0258
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-448-196A-3

  Query Match             100.0%;  Score 123;  DB 1;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 4
US-08-984-176-1
; Sequence 1, Application US/08984176
; Patent No. 5948609
; GENERAL INFORMATION:
;  APPLICANT: CARTER, DANIEL C
;  APPLICANT: HO, JOSEPH X
;  APPLICANT: RUKER, FLORIAN
;  TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
;  TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER
;  FILE REFERENCE: 08/984,176
;  CURRENT APPLICATION NUMBER: US/08/984,176
;  CURRENT FILING DATE: 1997-12-03
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 585
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-984-176-1

  Query Match             100.0%;  Score 123;  DB 2;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 5
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
;  GENERAL INFORMATION:
;    APPLICANT:  Kerry-Williams, Sean M
;    APPLICANT:  Gilbert, Sarah C
;    TITLE OF INVENTION:  Yeast Strains and Modified Albumins
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Centeon L.L.C.
;      STREET:  1020 First Avenue
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406-1310
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  MicroSoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/702,572
;      FILING DATE:  11-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO 95/23857
;      FILING DATE:  1-MAR-1995
;      APPLICATION NUMBER:  GB 9404270.2
;      FILING DATE:  5-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Naomi Biswas
;      REGISTRATION NUMBER:  38,384
;      REFERENCE/DOCKET NUMBER:  CE0114 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610/878/4294
;      TELEFAX:  610/878/4221
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-702-572-2

  Query Match             100.0%;  Score 123;  DB 2;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 6
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
;  GENERAL INFORMATION:
;    APPLICANT:  Sonnenschein, Carlos
;    APPLICANT:  Soto, Ana M.
;    TITLE OF INVENTION:  Inhibiting Proliferation of Cancer Cells
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/769,746
;      FILING DATE:  19-DEC-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Carroll, Peter G.
;      REGISTRATION NUMBER:  32,837
;      REFERENCE/DOCKET NUMBER:  MBRI-02584
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  585 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-769-746-2

  Query Match             100.0%;  Score 123;  DB 3;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
;  GENERAL INFORMATION:
;    APPLICANT:  Lichenstein, Henri
;    APPLICANT:  Lyons, David
;    APPLICANT:  Wurfel, Mark
;    APPLICANT:  Wright, Samuel
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/222,619
;      FILING DATE:
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  609 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-222-619-3

  Query Match             100.0%;  Score 123;  DB 1;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
;  GENERAL INFORMATION:
;    APPLICANT:  Sreekrishna, Kotikanyadan
;    APPLICANT:  Barr, Kathryn A.
```

```
;    APPLICANT:  Brierley, Russell A.
;    APPLICANT:  Thill, Gregory P.
;    APPLICANT:  Tschopp, Juerg F.
;    TITLE OF INVENTION:  EXPRESSION OF HUMAN SERUM ALBUMIN IN
;    TITLE OF INVENTION:  PICHIA PASTORIS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/433,037
;      FILING DATE:  03-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9108Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  609 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-433-037-4

  Query Match             100.0%;  Score 123;  DB 1;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 9
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
;  APPLICANT:  Mary Ellen Digan
;  APPLICANT:  Philip Lake
;  APPLICANT:  Hermann Gram
;  TITLE OF INVENTION:  Fusion Polypeptides
;  FILE REFERENCE:  600-7244/CPA
;  CURRENT APPLICATION NUMBER: US/08/897,956A
;  CURRENT FILING DATE:  1997-07-21
;  PRIOR APPLICATION NUMBER:  60/022,689
;  PRIOR FILING DATE:  1996-07-26
;  NUMBER OF SEQ ID NOS:  38
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 609
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-08-897-956A-2

  Query Match             100.0%;  Score 123;  DB 4;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 10
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 977
;   LENGTH: 609
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

  Query Match             100.0%;  Score 123;  DB 4;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 11
US-09-919-039-370
; Sequence 370, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
;  APPLICANT: Kaser, Matthew R.
;  TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
;  FILE REFERENCE: PA-0035 US
;  CURRENT APPLICATION NUMBER: US/09/919,039
;  CURRENT FILING DATE:  2002-09-09
;  PRIOR APPLICATION NUMBER: 60/222,113
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 401
;  SOFTWARE: PERL Program
; SEQ ID NO 370
;   LENGTH: 609
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6727066 088957CD1
US-09-919-039-370

  Query Match             100.0%;  Score 123;  DB 4;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 12
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  AMGEN INC.
;    TITLE OF INVENTION:  Afamin: A Human Serum Albumin-Like
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Center, Patent Operations/RRC
;      STREET:  1840 DeHavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04075
;      FILING DATE:
;      CLASSIFICATION:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  609 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
PCT-US95-04075-3

  Query Match             100.0%;  Score 123;  DB 5;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 13
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
;  GENERAL INFORMATION:
;    APPLICANT:  Fleer, Reinhard
;    APPLICANT:  Fournier, Alain
;    APPLICANT:  Guitton, Jean-Dominique
;    APPLICANT:  Jung, Gerard
;    APPLICANT:  Yeh, Patrice
;    TITLE OF INVENTION:  NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;    TITLE OF INVENTION:  PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;    TITLE OF INVENTION:  CONTAINING SAID POLYPEPTIDES
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  System 7.1
;      SOFTWARE:  Word 5.1 (PatentIn)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/797,689
;      FILING DATE:  31-JAN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/256,927
;      FILING DATE:  28-JUL-1994
;      APPLICATION NUMBER:  FR 92/01064
;      FILING DATE:  31-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00085
;      FILING DATE:  28-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith Ph.D., Julie K.
;      REGISTRATION NUMBER:  P-38,619
;      REFERENCE/DOCKET NUMBER:  ST92006-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3839
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  610 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-797-689-2

  Query Match             100.0%;  Score 123;  DB 2;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 14
US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686179
;  GENERAL INFORMATION:
;       APPLICANT: Fleer, Reinhard
;                  Fournier, Alain
;                  Guitton, Jean-Dominique
;                  Jung, Gerard
;                  Yeh, Patrice
;       TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;                           PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;                           CONTAINING SAID POLYPEPTIDES
;       NUMBER OF SEQUENCES: 36
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Rhone-Poulenc Rorer Inc.
;            STREET: 500 Arcola Road, 3C43
;            CITY: Collegeville
;            STATE: PA
;            COUNTRY: USA
;            ZIP: 19426
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: Macintosh
;            OPERATING SYSTEM: System 7.1
;            SOFTWARE: Word 5.1 (PatentIn)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/984,186
;            FILING DATE: 29-Oct-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/797,689
;            FILING DATE: 31-JAN-1997
;            APPLICATION NUMBER: US 08/256,927
;            FILING DATE: 28-JUL-1994
;            APPLICATION NUMBER: FR 92/01064
;            FILING DATE: 31-JAN-1992
;            APPLICATION NUMBER: PCT/FR93/00085
;            FILING DATE: 28-JAN-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Smith Ph.D., Julie K.
;            REGISTRATION NUMBER: P-38,619
;            REFERENCE/DOCKET NUMBER: ST92006-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (610) 454-3839
```

```
;            TELEFAX: (610) 454-3808
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 610 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

  Query Match            100.0%;  Score 123;  DB 4;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 15
US-09-949-016-11170
; Sequence 11170, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11170
;   LENGTH: 622
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-11170

  Query Match            100.0%;  Score 123;  DB 4;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db         38 DAHKSEVAHRFKDLGEENFKALVL 61


RESULT 16
US-08-256-938-2
; Sequence 2, Application US/08256938
; Patent No. 5665863
;  GENERAL INFORMATION:
;    APPLICANT:  Yeh, Patrice
;    TITLE OF INVENTION:  NEW POLYPEPTIDES HAVING GRANULOCYTE
;    TITLE OF INVENTION:  COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
;    TITLE OF INVENTION:  PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  System 7.1
;      SOFTWARE:  Word 5.0 (PatentIn)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/256,938
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 92/01065
;      FILING DATE:  31-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Goodman, Rosanne
;      REGISTRATION NUMBER:  32,534
;      REFERENCE/DOCKET NUMBER:  ST92007-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3817
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  783 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-256-938-2

  Query Match            100.0%;  Score 123;  DB 1;  Length 783;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 17
US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
;  GENERAL INFORMATION:
;    APPLICANT:  Yeh, Patrice
;    TITLE OF INVENTION:  NEW POLYPEPTIDES HAVING GRANULOCYTE
;    TITLE OF INVENTION:  COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
;    TITLE OF INVENTION:  PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  System 7.1
;      SOFTWARE:  Word 5.0 (PatentIn)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/256,938
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 92/01065
;      FILING DATE:  31-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Goodman, Rosanne
;      REGISTRATION NUMBER:  32,534
;      REFERENCE/DOCKET NUMBER:  ST92007-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3817
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  787 amino acids
;      TYPE:  amino acid
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-256-938-4

  Query Match             100.0%;  Score 123;  DB 1;  Length 787;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db        203 DAHKSEVAHRFKDLGEENFKALVL 226


RESULT 18
US-08-797-689-16
; Sequence 16, Application US/08797689
; Patent No. 5876969
;  GENERAL INFORMATION:
;    APPLICANT:  Fleer, Reinhard
;    APPLICANT:  Fournier, Alain
;    APPLICANT:  Guitton, Jean-Dominique
;    APPLICANT:  Jung, Gerard
;    APPLICANT:  Yeh, Patrice
;    TITLE OF INVENTION:  NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;    TITLE OF INVENTION:  PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;    TITLE OF INVENTION:  CONTAINING SAID POLYPEPTIDES
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  System 7.1
;      SOFTWARE:  Word 5.1 (PatentIn)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/797,689
;      FILING DATE:  31-JAN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/256,927
;      FILING DATE:  28-JUL-1994
;      APPLICATION NUMBER:  FR 92/01064
;      FILING DATE:  31-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00085
;      FILING DATE:  28-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith Ph.D., Julie K.
;      REGISTRATION NUMBER:  P-38,619
;      REFERENCE/DOCKET NUMBER:  ST92006-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 454-3839
;      TELEFAX:  (610) 454-3808
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  787 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-797-689-16

  Query Match             100.0%;  Score 123;  DB 2;  Length 787;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db        203 DAHKSEVAHRFKDLGEENFKALVL 226


RESULT 19
US-09-984-186-16
; Sequence 16, Application US/09984186
; Patent No. 6686179
;  GENERAL INFORMATION:
;       APPLICANT: Fleer, Reinhard
;                  Fournier, Alain
;                  Guitton, Jean-Dominique
;                  Jung, Gerard
;                  Yeh, Patrice
;       TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;                           PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;                           CONTAINING SAID POLYPEPTIDES
;       NUMBER OF SEQUENCES: 36
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Rhone-Poulenc Rorer Inc.
;            STREET: 500 Arcola Road, 3C43
;            CITY: Collegeville
;            STATE: PA
;            COUNTRY: USA
;            ZIP: 19426
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: Macintosh
;            OPERATING SYSTEM: System 7.1
;            SOFTWARE: Word 5.1 (PatentIn)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/984,186
;            FILING DATE: 29-Oct-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/797,689
;            FILING DATE: 31-JAN-1997
;            APPLICATION NUMBER: US 08/256,927
;            FILING DATE: 28-JUL-1994
;            APPLICATION NUMBER: FR 92/01064
;            FILING DATE: 31-JAN-1992
;            APPLICATION NUMBER: PCT/FR93/00085
;            FILING DATE: 28-JAN-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Smith Ph.D., Julie K.
;            REGISTRATION NUMBER: P-38,619
;            REFERENCE/DOCKET NUMBER: ST92006-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (610) 454-3839
;            TELEFAX: (610) 454-3808
;  INFORMATION FOR SEQ ID NO: 16:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 787 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-984-186-16

  Query Match             100.0%;  Score 123;  DB 4;  Length 787;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db        203 DAHKSEVAHRFKDLGEENFKALVL 226


RESULT 20
US-08-897-956A-3
; Sequence 3, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
;  APPLICANT:  Mary Ellen Digan
;  APPLICANT:  Philip Lake
```

```
;  APPLICANT:  Hermann Gram
;  TITLE OF INVENTION:  Fusion Polypeptides
;  FILE REFERENCE:  600-7244/CPA
;  CURRENT APPLICATION NUMBER: US/08/897,956A
;  CURRENT FILING DATE:  1997-07-21
;  PRIOR APPLICATION NUMBER:  60/022,689
;  PRIOR FILING DATE:  1996-07-26
;  NUMBER OF SEQ ID NOS:  38
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 978
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion polypeptide
US-08-897-956A-3

  Query Match             100.0%;  Score 123;  DB 4;  Length 978;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db        212 DAHKSEVAHRFKDLGEENFKALVL 235


RESULT 21
US-08-134-638-1
; Sequence 1, Application US/08134638
; Patent No. 5473050
;  GENERAL INFORMATION:
;    APPLICANT:  Strand, Frederick T
;    TITLE OF INVENTION:  Denatured Bovine Serum Albumin Milk
;    TITLE OF INVENTION:  Products and Method Therefor
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Frederick T. Strand
;      STREET:  P.O. Box 64321
;      CITY:  Phoenix
;      STATE:  Arizona
;      COUNTRY:  USA
;      ZIP:  85082-4321
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 1.2 Mb storage
;      COMPUTER:  IBM PC
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/134,638
;      FILING DATE:  10/12/93
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  N/A
;      FILING DATE:  N/A
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiss, Harry M
;      REGISTRATION NUMBER:  19,497
;      REFERENCE/DOCKET NUMBER:  1795P1423
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (602) 994-8888
;      TELEFAX:  (602) 947-2663
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  582
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-134-638-1

  Query Match             88.6%;  Score 109;  DB 1;  Length 582;
  Best Local Similarity   83.3%;  Pred. No. 3.3e-09;
  Matches   20;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||:||||||||||:|| |||
Db          1 DTHKSEIAHRFKDLGEEHFKGLVL 24


RESULT 22
US-08-448-196A-4
; Sequence 4, Application US/08448196A
; Patent No. 5780594
;  GENERAL INFORMATION:
;    APPLICANT:  CARTER, DANIEL C.
;    TITLE OF INVENTION:  BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
;    TITLE OF INVENTION:  CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
;    TITLE OF INVENTION:  RELATED PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NASA
;      STREET:  MARSHALL SPACE FLIGHT CENTER
;      CITY:  HUNTSVILLE
;      STATE:  ALABAMA
;      COUNTRY:  USA
;      ZIP:  35812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,196A
;      FILING DATE:  23-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROAD JR., ROBERT L.
;      REGISTRATION NUMBER:  18,757
;      REFERENCE/DOCKET NUMBER:  XX/MFS-28402-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  205-544-0021
;      TELEFAX:  205-544-0258
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  583 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-448-196A-4

  Query Match             88.6%;  Score 109;  DB 1;  Length 583;
  Best Local Similarity   83.3%;  Pred. No. 3.3e-09;
  Matches   20;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||:||||||||||:|| |||
Db          1 DTHKSEIAHRFKDLGEEHFKGLVL 24


RESULT 23
US-08-448-196A-7
; Sequence 7, Application US/08448196A
; Patent No. 5780594
;  GENERAL INFORMATION:
;    APPLICANT:  CARTER, DANIEL C.
;    TITLE OF INVENTION:  BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
;    TITLE OF INVENTION:  CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
;    TITLE OF INVENTION:  RELATED PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NASA
;      STREET:  MARSHALL SPACE FLIGHT CENTER
;      CITY:  HUNTSVILLE
;      STATE:  ALABAMA
```

```
;       COUNTRY:  USA
;       ZIP:  35812
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/448,196A
;       FILING DATE:  23-MAY-1995
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BROAD JR., ROBERT L.
;       REGISTRATION NUMBER:  18,757
;       REFERENCE/DOCKET NUMBER:  XX/MFS-28402-2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  205-544-0021
;       TELEFAX:  205-544-0258
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  584 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FRAGMENT TYPE:  N-terminal
US-08-448-196A-7

  Query Match             86.2%;  Score 106;  DB 1;  Length 584;
  Best Local Similarity   79.2%;  Pred. No. 9.9e-09;
  Matches   19;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :|||||:|||||||||::|| |||
Db          1 EAHKSEIAHRFKDLGEQHFKGLVL 24


RESULT 24
US-08-448-196A-6
; Sequence 6, Application US/08448196A
; Patent No. 5780594
;  GENERAL INFORMATION:
;    APPLICANT:  CARTER, DANIEL C.
;    TITLE OF INVENTION:  BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
;    TITLE OF INVENTION:  CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
;    TITLE OF INVENTION:  RELATED PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NASA
;      STREET:  MARSHALL SPACE FLIGHT CENTER
;      CITY:  HUNTSVILLE
;      STATE:  ALABAMA
;      COUNTRY:  USA
;      ZIP:  35812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,196A
;      FILING DATE:  23-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROAD JR., ROBERT L.
;      REGISTRATION NUMBER:  18,757
;      REFERENCE/DOCKET NUMBER:  XX/MFS-28402-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  205-544-0021
;      TELEFAX:  205-544-0258
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  583 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-448-196A-6

  Query Match             85.4%;  Score 105;  DB 1;  Length 583;
  Best Local Similarity   79.2%;  Pred. No. 1.4e-08;
  Matches   19;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||:|||| |||||||: |||
Db          1 DTHKSEIAHRFNDLGEENFQGLVL 24


RESULT 25
US-08-448-196A-5
; Sequence 5, Application US/08448196A
; Patent No. 5780594
;  GENERAL INFORMATION:
;    APPLICANT:  CARTER, DANIEL C.
;    TITLE OF INVENTION:  BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
;    TITLE OF INVENTION:  CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
;    TITLE OF INVENTION:  RELATED PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NASA
;      STREET:  MARSHALL SPACE FLIGHT CENTER
;      CITY:  HUNTSVILLE
;      STATE:  ALABAMA
;      COUNTRY:  USA
;      ZIP:  35812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,196A
;      FILING DATE:  23-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROAD JR., ROBERT L.
;      REGISTRATION NUMBER:  18,757
;      REFERENCE/DOCKET NUMBER:  XX/MFS-28402-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  205-544-0021
;      TELEFAX:  205-544-0258
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  583 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-448-196A-5

  Query Match             81.3%;  Score 100;  DB 1;  Length 583;
  Best Local Similarity   75.0%;  Pred. No. 9.1e-08;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||:|||| |||||::|| |||
Db          1 DTHKSEIAHRFNDLGEKHFKGLVL 24


RESULT 26
US-09-845-764A-1
```

```
; Sequence 1, Application US/09845764A
; Patent No. 6693080
; GENERAL INFORMATION:
;  APPLICANT: Jackowski, George
;  TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
;  TITLE OF INVENTION:  OF 1521 DALTONS
;  FILE REFERENCE: 2132.037
;  CURRENT APPLICATION NUMBER: US/09/845,764A
;  CURRENT FILING DATE:  2001-04-30
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-845-764A-1

  Query Match             60.2%;  Score 74;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 2e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDL 14
              ||||||||||||||
Db          2 DAHKSEVAHRFKDL 15


RESULT 27
US-08-470-187-1
; Sequence 1, Application US/08470187
; Patent No. 5532152
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine E.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor Acetyl
;    TITLE OF INVENTION:  Hydrolase
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gertein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,187
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5532152and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  31672
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-470-187-1

  Query Match             56.9%;  Score 70;  DB 1;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FKDLGEENFKALVL 24
              ||||||||||||||
Db          1 FKDLGEENFKALVL 14


RESULT 28
US-08-318-905-1
; Sequence 1, Application US/08318905
; Patent No. 5641669
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor Acetyl
;    TITLE OF INVENTION:  Hydrolase
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gertein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/318,905
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  6-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5641669and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  32205
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-318-905-1

  Query Match             56.9%;  Score 70;  DB 1;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FKDLGEENFKALVL 24
              ||||||||||||||
Db          1 FKDLGEENFKALVL 14


RESULT 29
US-08-483-232-1
; Sequence 1, Application US/08483232
; Patent No. 5656431
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,232
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5656431and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32689
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-483-232-1

  Query Match             56.9%;  Score 70;  DB 1;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FKDLGEENFKALVL 24
              ||||||||||||||
Db          1 FKDLGEENFKALVL 14


RESULT 30
US-08-483-140-1
; Sequence 1, Application US/08483140
; Patent No. 5698403
;  GENERAL INFORMATION:
;    APPLICANT:  ICOS Corporation
;    TITLE OF INVENTION:  Platelet-Activating Factor Acetyl
;    TITLE OF INVENTION:  Hydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gertein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,140
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  6-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  6-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5698403and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  32781
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-483-140-1

  Query Match             56.9%;  Score 70;  DB 1;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FKDLGEENFKALVL 24
              ||||||||||||||
Db          1 FKDLGEENFKALVL 14


RESULT 31
US-08-485-938A-1
; Sequence 1, Application US/08485938A
; Patent No. 5847088
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,938A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5847088and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32792
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-485-938A-1

  Query Match             56.9%;  Score 70;  DB 2;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FKDLGEENFKALVL 24
              ||||||||||||||
Db          1 FKDLGEENFKALVL 14


RESULT 32
US-08-910-041-1
; Sequence 1, Application US/08910041
; Patent No. 5977308
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/910,041
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/483,232
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/34026
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-910-041-1

  Query Match             56.9%;  Score 70;  DB 2;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FKDLGEENFKALVL 24
              ||||||||||||||
Db          1 FKDLGEENFKALVL 14


RESULT 33
US-09-328-474-1
; Sequence 1, Application US/09328474
; Patent No. 6045794
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/328,474
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/483,232
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/34026
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-328-474-1
```

```
  Query Match           56.9%;  Score 70;  DB 3;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FKDLGEENFKALVL 24
              ||||||||||||||
Db          1 FKDLGEENFKALVL 14


RESULT 34
US-09-100-546-1
; Sequence 1, Application US/09100546
; Patent No. 6099836
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,546
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/010,715
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6099836and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32793
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-546-1

  Query Match           56.9%;  Score 70;  DB 3;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FKDLGEENFKALVL 24
              ||||||||||||||
Db          1 FKDLGEENFKALVL 14


RESULT 35
US-09-010-715-1
; Sequence 1, Application US/09010715
; Patent No. 6146625
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/010,715
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/318,905
;      FILING DATE:  06-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6146625and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32793
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-010-715-1

  Query Match           56.9%;  Score 70;  DB 3;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FKDLGEENFKALVL 24
              ||||||||||||||
Db          1 FKDLGEENFKALVL 14


RESULT 36
US-09-577-758-1
; Sequence 1, Application US/09577758
; Patent No. 6203790
;  GENERAL INFORMATION:
;    APPLICANT:  Cousens, Lawrence S.
;    APPLICANT:  Eberhardt, Christine D.
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Le Trong, Hai
;    APPLICANT:  Tjoelker, Larry W.
;    APPLICANT:  Wilder, Cheryl L.
;    TITLE OF INVENTION:  Platelet-Activating Factor
;    TITLE OF INVENTION:  Acetylhydrolase
```

```
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/577,758
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/010,715
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/133,803
;      FILING DATE:  06-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6203790and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/32793
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3658
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-577-758-1

  Query Match             56.9%;  Score 70;  DB 3;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FKDLGEENFKALVL 24
              ||||||||||||||
Db          1 FKDLGEENFKALVL 14


RESULT 37
US-08-803-364-7
; Sequence 7, Application US/08803364
; Patent No. 5864014
;  GENERAL INFORMATION:
;    APPLICANT:  FASANO, Alessio
;    TITLE OF INVENTION:  ZONULA OCCLUDENS TOXIN RECEPTOR
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W., Suite 800
;      CITY:  Washington, D.C.
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/803,364
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/___,___
;      FILING DATE:  20 FEB 1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KIT, Gordon
;      REGISTRATION NUMBER:  30,764
;      REFERENCE/DOCKET NUMBER:  A-6888
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
US-08-803-364-7

  Query Match             55.3%;  Score 68;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AHKSEVAHRFKDL 14
              |||||||||||||
Db          1 AHKSEVAHRFKDL 13


RESULT 38
US-09-024-198-13
; Sequence 13, Application US/09024198
; Patent No. 5912323
;  GENERAL INFORMATION:
;    APPLICANT:  FASANO, Alessio
;    TITLE OF INVENTION:  ZONULA OCCLUDENS TOXIN RECEPTOR
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W., Suite 800
;      CITY:  Washington, D.C.
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/024,198
;      FILING DATE:  17 FEB 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/803,364
;      FILING DATE:  20 FEB 1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KIT, Gordon
;      REGISTRATION NUMBER:  30,764
;      REFERENCE/DOCKET NUMBER:  A-6988
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
US-09-024-198-13

  Query Match             55.3%;  Score 68;  DB 2;  Length 13;
```

```
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AHKSEVAHRFKDL 14
              |||||||||||||
Db          1 AHKSEVAHRFKDL 13


RESULT 39
US-09-186-409-13
; Sequence 13, Application US/09186409
; Patent No. 5948629
;  GENERAL INFORMATION:
;    APPLICANT:  FASANO, Alessio
;    TITLE OF INVENTION:  ZONULA OCCLUDENS TOXIN RECEPTOR
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W., Suite 800
;      CITY:  Washington, D.C.
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/186,409
;      FILING DATE:  5 NOV 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/024,198
;      FILING DATE:  17 FEB 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/803,364
;      FILING DATE:  20 FEB 1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KIT, Gordon
;      REGISTRATION NUMBER:  30,764
;      REFERENCE/DOCKET NUMBER:  A-7306
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
US-09-186-409-13

  Query Match             55.3%;  Score 68;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AHKSEVAHRFKDL 14
              |||||||||||||
Db          1 AHKSEVAHRFKDL 13


RESULT 40
US-08-803-364-1
; Sequence 1, Application US/08803364
; Patent No. 5864014
;  GENERAL INFORMATION:
;    APPLICANT:  FASANO, Alessio
;    TITLE OF INVENTION:  ZONULA OCCLUDENS TOXIN RECEPTOR
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W., Suite 800
;      CITY:  Washington, D.C.
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/803,364
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/___,___
;      FILING DATE:  20 FEB 1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KIT, Gordon
;      REGISTRATION NUMBER:  30,764
;      REFERENCE/DOCKET NUMBER:  A-6888
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
US-08-803-364-1

  Query Match             51.2%;  Score 63;  DB 2;  Length 16;
  Best Local Similarity   85.7%;  Pred. No. 0.0012;
  Matches   12;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGE 16
              |||| ||||||| |
Db          3 HKSEAAHRFKDLQE 16


Search completed: August 19, 2005, 10:59:14
Job time : 24 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 10:48:50 ; Search time 163 Seconds
(without alignments)
56.946 Million cell updates/sec

Title: US-09-846-328B-1_COPY_2_25
Perfect score: 123
Sequence: 1 DAHKSEVAHRFKDLGEENFKALVL 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	123	100.0	26	6	ABU08846	Abu08846 Serum alb
2	123	100.0	26	6	ABG75942	Abg75942 Human ser
3	123	100.0	28	6	ABU08847	Abu08847 Serum alb
4	123	100.0	28	6	ABU09223	Abu09223 Serum alb
5	123	100.0	116	4	AAO02642	Aao02642 Human pol
6	123	100.0	119	4	AAO11632	Aao11632 Human pol
7	123	100.0	133	1	AAP30004	Aap30004 Sequence
8	123	100.0	192	4	AAU29875	Aau29875 Novel hum
9	123	100.0	195	5	AAO17048	Aao17048 Human ser
10	123	100.0	195	7	ABU10022	Abu10022 Human ser
11	123	100.0	204	3	AAY83947	Aay83947 Yeast cod
12	123	100.0	214	4	AAU29874	Aau29874 Novel hum
13	123	100.0	236	5	AAO17051	Aao17051 Human alb
14	123	100.0	241	5	AAO16984	Aao16984 Alpha-MSH
15	123	100.0	241	7	ABU10025	Abu10025 Alpha-MSH
16	123	100.0	242	5	AAO16985	Aao16985 Alpha-MSH
17	123	100.0	244	5	AAO16986	Aao16986 Alpha-MSH
18	123	100.0	245	5	AAO16987	Aao16987 Alpha-MSH
19	123	100.0	245	5	AAO16988	Aao16988 Alpha-MSH
20	123	100.0	268	5	AAO16989	Aao16989 Alpha-MSH
21	123	100.0	268	7	ABU10026	Abu10026 Alpha-MSH
22	123	100.0	289	4	AAU29581	Aau29581 Novel hum
23	123	100.0	289	4	AAU29575	Aau29575 Novel hum
24	123	100.0	303	2	AAR14178	Aar14178 Human ser
25	123	100.0	373	1	AAP90387	Aap90387 N-termina
26	123	100.0	388	1	AAP90389	Aap90389 N-termina
27	123	100.0	389	1	AAP90390	Aap90390 N-termina
28	123	100.0	390	1	AAP90391	Aap90391 N-termina
29	123	100.0	401	4	AAU29876	Aau29876 Novel hum
30	123	100.0	407	1	AAP90392	Aap90392 N-termina
31	123	100.0	500	7	ADD32019	Add32019 Heterolog
32	123	100.0	544	8	ABM84998	Abm84998 Human dia
33	123	100.0	550	4	AAU29877	Aau29877 Novel hum
34	123	100.0	571	8	ABM84997	Abm84997 Human dia
35	123	100.0	573	8	ABM84996	Abm84996 Human dia
36	123	100.0	573	8	ABM84995	Abm84995 Human dia
37	123	100.0	584	6	ABG72381	Abg72381 Mature hu
38	123	100.0	585	1	AAP93344	Aap93344 Sequence
39	123	100.0	585	1	AAP90388	Aap90388 Mature hu
40	123	100.0	585	1	AAP91422	Aap91422 Human nor
41	123	100.0	585	2	AAR05318	Aar05318 Human ser
42	123	100.0	585	2	AAR08457	Aar08457 Human ser
43	123	100.0	585	2	AAR26207	Aar26207 Human ser
44	123	100.0	585	2	AAR26362	Aar26362 Synthetic
45	123	100.0	585	2	AAR20029	Aar20029 Human ser

ALIGNMENTS

RESULT 1
ABU08846
ID ABU08846 standard; peptide; 26 AA.
XX
AC ABU08846;
XX
DT 25-AUG-2003 (first entry)
XX
DE Serum albumin peptide, #3, used for physiologial condition diagnostics.
XX
KW Proteomic; human; physiological condition; analyte; biopolymer;
KW biomarker; serum albumin; insulin resistance.
XX
OS Homo sapiens.
XX
PN US2002160420-A1.
XX
PD 31-OCT-2002.
XX
PF 30-APR-2001; 2001US-00846330.
XX
PR 30-APR-2001; 2001US-00846330.
XX
PA (JACK/) JACKOWSKI G.
PA (THAT/) THATCHER B.
PA (MARS/) MARSHALL J.
PA (YANT/) YANTHA J.
PA (VREE/) VREES T.
XX
PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;
XX
DR WPI; 2003-491923/46.
XX
PT Determining proteomic basis e.g. basis for diagnosing existence of or
PT predicting development and/or progression of abnormal physiological
PT conditions based upon the presence of proteomic materials.
XX
PS Disclosure; Page 22; 25pp; English.
XX
CC The invention discloses a method for determining a proteomic basis for
CC development and progression of abnormal physiological conditions. The
CC method comprises isolating one or more patient specific proteomic
CC materials from a sample and comparing it against a library of proteomic
CC materials having characteristics identifiable with both normal and
CC abnormal physiological conditions or their predictive hallmarks. The
CC method is useful for determining a proteomic basis for development and
CC progression of abnormal physiological conditions. The method is also
CC useful for evaluating samples containing several analytes/biopolymers for

```
CC   the presence of physiological condition specific sequences. The peptide
CC   presented is a biomarker from serum albumin and is associated with
CC   insulin resistance
XX
SQ   Sequence 26 AA;

  Query Match             100.0%;  Score 123;  DB 6;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-12;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db          2 DAHKSEVAHRFKDLGEENFKALVL 25


RESULT 2
ABG75942
ID   ABG75942 standard; peptide; 26 AA.
XX
AC   ABG75942;
XX
DT   08-MAY-2003  (first entry)
XX
DE   Human serum albumin biopolymer marker peptide.
XX
KW   Human; serum albumin; biopolymer marker; SELDI;
KW   Surface Enhanced Laser Desorption Ionization mass spectrometry;
KW   time-of-flight detection procedure; complement system disease;
KW   syndrome X; insulin resistance; hyperinsulinaemia.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Optionally absent"
FT   Misc-difference 2. .25
FT                   /note= "Specifically claimed in claim 1"
FT   Misc-difference 26
FT                   /note= "Optionally absent"
XX
PN   US2002160531-A1.
XX
PD   31-OCT-2002.
XX
PF   30-APR-2001; 2001US-00846328.
XX
PR   30-APR-2001; 2001US-00846328.
XX
PA   (JACK/) JACKOWSKI G.
PA   (THAT/) THATCHER B.
PA   (MARS/) MARSHALL J.
PA   (YANT/) YANTHA J.
PA   (VREE/) VREES T.
XX
PI   Jackowski G,  Thatcher B,  Marshall J,  Yantha J,  Vrees T;
XX
DR   WPI; 2003-255195/25.
XX
PT   A new biopolymer marker useful in indicating a particular disease state
PT   particularly insulin resistance.
XX
PS   Disclosure; Fig 1; 10pp; English.
XX
CC   The invention describes a biopolymer marker useful in indicating at least
CC   one particular disease state. Biopolymer markers are identified using
CC   Surface Enhanced Laser Desorption Ionization (SELDI) mass spectrometry
CC   and time-of-flight detection procedures. The markers are useful for
CC   indicating e.g. complement system diseases and syndrome X, characterised
CC   by the clustering of insulin resistance and hyperinsulinaemia. This is
CC   the amino acid sequence of a human serum albumin biopolymer marker useful
CC   for indicating insulin resistance
XX
SQ   Sequence 26 AA;

  Query Match             100.0%;  Score 123;  DB 6;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-12;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db          2 DAHKSEVAHRFKDLGEENFKALVL 25


RESULT 3
ABU08847
ID   ABU08847 standard; peptide; 28 AA.
XX
AC   ABU08847;
XX
DT   25-AUG-2003  (first entry)
XX
DE   Serum albumin peptide, #4, used for physiologial condition diagnostics.
XX
KW   Proteomic; human; physiological condition; analyte; biopolymer;
KW   biomarker; serum albumin; insulin resistance.
XX
OS   Homo sapiens.
XX
PN   US2002160420-A1.
XX
PD   31-OCT-2002.
XX
PF   30-APR-2001; 2001US-00846330.
XX
PR   30-APR-2001; 2001US-00846330.
XX
PA   (JACK/) JACKOWSKI G.
PA   (THAT/) THATCHER B.
PA   (MARS/) MARSHALL J.
PA   (YANT/) YANTHA J.
PA   (VREE/) VREES T.
XX
PI   Jackowski G,  Thatcher B,  Marshall J,  Yantha J,  Vrees T;
XX
DR   WPI; 2003-491923/46.
XX
PT   Determining proteomic basis e.g. basis for diagnosing existence of or
PT   predicting development and/or progression of abnormal physiological
PT   conditions based upon the presence of proteomic materials.
XX
PS   Disclosure; Page 22; 25pp; English.
XX
CC   The invention discloses a method for determining a proteomic basis for
CC   development and progression of abnormal physiological conditions. The
CC   method comprises isolating one or more patient specific proteomic
CC   materials from a sample and comparing it against a library of proteomic
CC   materials having characteristics identifiable with both normal and
CC   abnormal physiological conditions or their predictive hallmarks. The
CC   method is useful for determining a proteomic basis for development and
CC   progression of abnormal physiological conditions. The method is also
CC   useful for evaluating samples containing several analytes/biopolymers for
CC   the presence of physiological condition specific sequences. The peptide
CC   presented is a biomarker from serum albumin and is associated with
CC   insulin resistance
XX
SQ   Sequence 28 AA;

  Query Match             100.0%;  Score 123;  DB 6;  Length 28;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-12;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db          2 DAHKSEVAHRFKDLGEENFKALVL 25
```

```
RESULT 4
ABU09223
ID   ABU09223 standard; peptide; 28 AA.
XX
AC   ABU09223;
XX
DT   13-JUN-2003  (first entry)
XX
DE   Serum albumin insulin resistance disease specific marker.
XX
KW   Human; insulin resistance; serum albumin; biopolymer marker; lipaemia;
KW   complement system disease; syndrome X; obesity; hyperinsulinaemia;
KW   glucose intolerance; blood coagulation abnormality; hyperuricaemia;
KW   albuminuria; kidney failure; stroke; hypertension; diabetes;
KW   heart failure.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          2. .27
FT                   /note= "Specifically claimed in claim 1"
XX
PN   US2002161177-A1.
XX
PD   31-OCT-2002.
XX
PF   30-APR-2001; 2001US-00846329.
XX
PR   30-APR-2001; 2001US-00846329.
XX
PA   (JACK/) JACKOWSKI G.
PA   (THAT/) THATCHER B.
PA   (MARS/) MARSHALL J.
PA   (YANT/) YANTHA J.
PA   (VREE/) VREES T.
XX
PI   Jackowski G,  Thatcher B,  Marshall J,  Yantha J,  Vrees T;
XX
DR   WPI; 2003-361743/34.
XX
PT   Serum albumin biopolymer marker having specific molecular weight
PT   indicates at least one particular disease state.
XX
PS   Disclosure; Fig 1; 10pp; English.
XX
CC   The invention relates to a serum albumin biopolymer marker having a
CC   molecular weight of 2937 daltons which indicates at least one particular
CC   disease state. The bioplymer marker is useful for indicating at least
CC   one particular disease state, e.g. insulin resistance. Also for diseases
CC   associated with the complement system and syndrome X such as obesity,
CC   hyperinsulinaemia, lipaemia, glucose intolerance, blood coagulation
CC   abnormalities, hyperuricaemia, albuminuria, kidney failure, stroke,
CC   hypertension, diabetes and heart failure. The biopolymer marker enables a
CC   diagnostician to gain the ability to characterise either the presence or
CC   absence of the disease state relative to recognition of the presence
CC   and/or the absence of the biopolymer. The present sequence represents the
CC   amino acid sequence of the serum albumin insulin resistance disease
CC   specific marker
XX
SQ   Sequence 28 AA;

  Query Match             100.0%;  Score 123;  DB 6;  Length 28;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-12;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db          2 DAHKSEVAHRFKDLGEENFKALVL 25


RESULT 5
AAO02642
ID   AAO02642 standard; protein; 116 AA.
XX
AC   AAO02642;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 16534.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US004927.
XX
PR   28-FEB-2000; 2000US-00515126.
PR   18-MAY-2000; 2000US-00577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   N-PSDB; AAI82573.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT   and treating e.g. leukemia, inflammation and immune disorders.
XX
PS   Claim 20; SEQ ID NO 16534; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 116 AA;

  Query Match             100.0%;  Score 123;  DB 4;  Length 116;
  Best Local Similarity   100.0%;  Pred. No. 8.6e-12;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db         22 DAHKSEVAHRFKDLGEENFKALVL 45


RESULT 6
AAO11632
ID   AAO11632 standard; protein; 119 AA.
XX
AC   AAO11632;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 25524.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
```

```
KW   nervous system disorders; arthritis; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US004927.
XX
PR   28-FEB-2000; 2000US-00515126.
PR   18-MAY-2000; 2000US-00577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   N-PSDB; AAI91563.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT   and treating e.g. leukemia, inflammation and immune disorders.
XX
PS   Claim 20; SEQ ID NO 25524; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 119 AA;

  Query Match             100.0%;  Score 123;  DB 4;  Length 119;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-12;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         30 DAHKSEVAHRFKDLGEENFKALVL 53


RESULT 7
AAP30004
ID   AAP30004 standard; protein; 133 AA.
XX
AC   AAP30004;
XX
DT   25-APR-1992  (first entry)
XX
DE   Sequence encoded by recombinant DNA molecule pKT218(HSA/33-1 (Bg1II-
DE   EcoRI)-HSA/17-3(Bg1II-EcoRI)) human proserum albumin.
XX
KW   Burn therapy; blood; protein; nitrogen fixation; food additive;
KW   kernicterus therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .26
FT                   /label= signal
FT   Protein         27. .59
FT                   /label= residues 1-33
FT   Protein         60. .92
FT                   /label= residues 319-351
FT   Protein         93. .124
FT                   /label= residues 368-399
FT   Protein         125. .133
FT                   /label= residues 572-580
XX
PN   EP91527-A.
XX
PD   19-OCT-1983.
XX
PF   14-DEC-1981;   81US-00330912.
XX
PR   14-DEC-1981;   81US-00330912.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Gilbert W,  Philipp BW;
XX
DR   WPI; 1983-796520/43.
DR   N-PSDB; AAN30009.
XX
PT   Prodn. of human serum albumin-like polypeptide(s) - from host transformed
PT   with recombinant DNA molecule.
XX
PS   Example; Fig 4; 42pp; English.
XX
CC   AAN30009 was prepd. from two clones isolated from a human fetal liver
CC   cDNA library screened with mouse serum albumin cDNA. The two clones are
CC   pKT218(HSA/33-1) and pKT218(HSA/17-3), which were recombined to give
CC   AAN30009, the coding sequence of which is out of phase by a single
CC   nucleoticde (see FT). However, as a result of perhaps some internal
CC   start, hosts transformed with the hybrid gene (even out of phase) in
CC   pKT218 still produce HSA-like products
XX
SQ   Sequence 133 AA;

  Query Match             100.0%;  Score 123;  DB 1;  Length 133;
  Best Local Similarity   100.0%;  Pred. No. 1e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         27 DAHKSEVAHRFKDLGEENFKALVL 50


RESULT 8
AAU29875
ID   AAU29875 standard; protein; 192 AA.
XX
AC   AAU29875;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secreted protein #366.
XX
KW   Human; vaccination; gene therapy; nutritional supplement;
KW   stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW   immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200179449-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US008656.
XX
PR   18-APR-2000; 2000US-00552929.
PR   26-JAN-2001; 2001US-00770160.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
```

DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 206; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 192 AA;

Query Match 100.0%; Score 123; DB 4; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 53 DAHKSEVAHRFKDLGEENFKALVL 76

RESULT 9
AAO17048
ID AAO17048 standard; protein; 195 AA.
XX
AC AAO17048;
XX
DT 29-MAY-2002 (first entry)
XX
DE Human serum albumin (1-195) SEQ ID NO: 56.
XX
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX
OS Homo sapiens.
XX
PN WO200206316-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US022263.
XX
PR 14-JUL-2000; 2000US-0218381P.
PR 18-AUG-2000; 2000US-0226382P.
PR 06-OCT-2000; 2000US-0238380P.
PR 29-DEC-2000; 2000US-0258764P.
PR 14-JUN-2001; 2001US-0298317P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX
DR WPI; 2002-195801/25.
XX
PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX
PS Example 2; Page 46; 89pp; English.
XX
CC The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention
XX
SQ Sequence 195 AA;

Query Match 100.0%; Score 123; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.6e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 1 DAHKSEVAHRFKDLGEENFKALVL 24

RESULT 10
ABU10022
ID ABU10022 standard; protein; 195 AA.
XX
AC ABU10022;
XX
DT 31-JUL-2003 (first entry)
XX
DE Human serum albumin residues 1-195.
XX
KW Bladder disorder; cytostatic; antiinflammatory; immune response;
KW un-methylated CpG sequence; alpha-MSH; melanocortin receptor;
KW bladder cancer; tumour; interstitial cystitis; inflammation;
KW alpha-MSH concatamer; melanocyte stimulating hormone; human;
KW serum albumin.
XX
OS Homo sapiens.
XX
PN US2002193332-A1.
XX
PD 19-DEC-2002.
XX
PF 12-FEB-2002; 2002US-00074956.
XX
PR 12-FEB-2001; 2001US-0268175P.
XX
PA (HEDL/) HEDLEY M L.
XX
PI Hedley ML;
XX
DR WPI; 2003-447327/42.
XX
PT Modulating immune responses in a mammal with a bladder disorder e.g.
PT bladder cancer, by administering nucleic acids comprising un-methylated
PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to
PT the mammal.
XX
PS Example 2; Page 9; 17pp; English.
XX
CC The invention describes a method of modulating an immune response in a
CC mammal, comprising identifying a mammal that has or is at risk for having
CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)
CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated
CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;

CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The
CC method is useful for modulating immune response in a mammal having a
CC bladder disorder, where administration of (N1) results in an amelioration
CC of one or more symptoms of the disorder. Preferably, the method is useful
CC for modulating immune response in a mammal having bladder cancer (where
CC administration of (N1) results in a decrease in tumour size or activity),
CC or for modulating immune response in a mammal having interstitial
CC cystitis (where administration of (N1) results in a modulation of the
CC immune response from Th2 response to a Th1 response). The method is also
CC useful for modulating immune response in a mammal having bladder disorder
CC that is characterised by inflammation which is associated with symptoms
CC of interstitial cystitis or associated with a disruption of the integrity
CC of the bladder lining. This is the amino acid sequence of human serum
CC albumin residues 1-195 that can be used in the creation of melanocyte
CC stimulating hormone (alpha-MSH) concatamers resulting in secretion of the
CC fusion protein when expressed in mammalian cells
XX
SQ Sequence 195 AA;

Query Match 100.0%; Score 123; DB 7; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.6e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 1 DAHKSEVAHRFKDLGEENFKALVL 24

RESULT 11
AAY83947
ID AAY83947 standard; protein; 204 AA.
XX
AC AAY83947;
XX
DT 28-JUL-2000 (first entry)
XX
DE Yeast codon-biased recombinant HSA protein fragment HSA-I.
XX
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
KW overlapping oligonucleotide; expression vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CN1239103-A.
XX
PD 22-DEC-1999.
XX
PF 17-JUN-1998; 98CN-00102506.
XX
PR 17-JUN-1998; 98CN-00102506.
XX
PA (HAIJ-) HAIJI BIOENGINEERING CO LTD.
XX
PI Li S, Lu D;
XX
DR WPI; 2000-351198/31.
DR N-PSDB; AAA10092.
XX
PT Process for preparing recombined human serum albumin comprising yeast
PT biased sex codons - uses a recombinant DNA technique.
XX
PS Example 1; Fig 3; 44pp; Chinese.
XX
CC The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as
CC three synthetic fragments (AAA10092-A10094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the sequence of the HSA fragment HSA-I encoded by the human gene with a
CC yeast codon bias. The invention also covers a recombinant expression
CC vector, yeast host cells carrying the recombinant expression vector and
CC the process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 123; DB 3; Length 204;
Best Local Similarity 100.0%; Pred. No. 1.6e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 8 DAHKSEVAHRFKDLGEENFKALVL 31

RESULT 12
AAU29874
ID AAU29874 standard; protein; 214 AA.
XX
AC AAU29874;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #365.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 205; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 214 AA;

Query Match 100.0%; Score 123; DB 4; Length 214;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         37 DAHKSEVAHRFKDLGEENFKALVL 60


RESULT 13
AAO17051
ID   AAO17051 standard; protein; 236 AA.
XX
AC   AAO17051;
XX
DT   29-MAY-2002  (first entry)
XX
DE   Human albumin-thrombin-alpha-MSH SEQ ID NO: 82.
XX
KW   Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW   alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW   cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW   immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW   antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW   antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW   diabetes; uveitis; coeliac disease.
XX
OS   Homo sapiens.
XX
PN   WO200206316-A2.
XX
PD   24-JAN-2002.
XX
PF   16-JUL-2001; 2001WO-US022263.
XX
PR   14-JUL-2000; 2000US-0218381P.
PR   18-AUG-2000; 2000US-0226382P.
PR   06-OCT-2000; 2000US-0238380P.
PR   29-DEC-2000; 2000US-0258764P.
PR   14-JUN-2001; 2001US-0298317P.
XX
PA   (ZYCO-) ZYCOS INC.
XX
PI   Hedley ML,  Urban R,  Aziz N,  Chen H,  Etemad-Moghadam B,  Yin P;
XX
DR   WPI; 2002-195801/25.
XX
PT   Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT   stimulating hormone concatamer or its analog, for treating inflammatory
PT   or autoimmune disorders.
XX
PS   Example 2; Page 48; 89pp; English.
XX
CC   The present invention relates to a nucleic acid comprising a sequence
CC   encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC   hormone (MSH) concatamer. The sequences are useful for treating an
CC   individual suffering from, or at risk of, a disorder of the immune system
CC   e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC   arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC   hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC   multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC   present sequence is a peptide described in the exemplification of the
CC   invention
XX
SQ   Sequence 236 AA;

  Query Match             100.0%;  Score 123;  DB 5;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 14
AAO16984
ID   AAO16984 standard; protein; 241 AA.
XX
AC   AAO16984;
XX
DT   29-MAY-2002  (first entry)
XX
DE   Alpha-MSH construct protein fragment SEQ ID NO: 59.
XX
KW   Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW   alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW   cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW   immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW   antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW   antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW   diabetes; uveitis; coeliac disease.
XX
OS   Unidentified.
XX
PN   WO200206316-A2.
XX
PD   24-JAN-2002.
XX
PF   16-JUL-2001; 2001WO-US022263.
XX
PR   14-JUL-2000; 2000US-0218381P.
PR   18-AUG-2000; 2000US-0226382P.
PR   06-OCT-2000; 2000US-0238380P.
PR   29-DEC-2000; 2000US-0258764P.
PR   14-JUN-2001; 2001US-0298317P.
XX
PA   (ZYCO-) ZYCOS INC.
XX
PI   Hedley ML,  Urban R,  Aziz N,  Chen H,  Etemad-Moghadam B,  Yin P;
XX
DR   WPI; 2002-195801/25.
XX
PT   Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT   stimulating hormone concatamer or its analog, for treating inflammatory
PT   or autoimmune disorders.
XX
PS   Example 2; Page 4-5; 89pp; English.
XX
CC   The present invention relates to a nucleic acid comprising a sequence
CC   encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC   hormone (MSH) concatamer. The sequences are useful for treating an
CC   individual suffering from, or at risk of, a disorder of the immune system
CC   e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC   arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC   hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC   multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC   present sequence is a protein described in the exemplification of the
CC   invention
XX
SQ   Sequence 241 AA;

  Query Match             100.0%;  Score 123;  DB 5;  Length 241;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 15
ABU10025
ID   ABU10025 standard; protein; 241 AA.
XX
AC   ABU10025;
XX
DT   31-JUL-2003  (first entry)
XX
```

DE Alpha-MSH/Serum albumin fusion protein H9.
XX
KW Bladder disorder; cytostatic; antiinflammatory; immune response;
KW un-methylated CpG sequence; alpha-MSH; melanocortin receptor;
KW bladder cancer; tumour; interstitial cystitis; inflammation;
KW alpha-MSH concatamer; melanocyte stimulating hormone; human;
KW serum albumin; fusion protein; H9.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2002193332-A1.
XX
PD 19-DEC-2002.
XX
PF 12-FEB-2002; 2002US-00074956.
XX
PR 12-FEB-2001; 2001US-0268175P.
XX
PA (HEDL/) HEDLEY M L.
XX
PI Hedley ML;
XX
DR WPI; 2003-447327/42.
XX
PT Modulating immune responses in a mammal with a bladder disorder e.g.
PT bladder cancer, by administering nucleic acids comprising un-methylated
PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to
PT the mammal.
XX
PS Example 2; Page 9; 17pp; English.
XX
CC The invention describes a method of modulating an immune response in a
CC mammal, comprising identifying a mammal that has or is at risk for having
CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)
CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated
CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;
CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The
CC method is useful for modulating immune response in a mammal having a
CC bladder disorder, where administration of (N1) results in an amelioration
CC of one or more symptoms of the disorder. Preferably, the method is useful
CC for modulating immune response in a mammal having bladder cancer (where
CC administration of (N1) results in a decrease in tumour size or activity),
CC or for modulating immune response in a mammal having interstitial
CC cystitis (where administration of (N1) results in a modulation of the
CC immune response from Th2 response to a Th1 response). The method is also
CC useful for modulating immune response in a mammal having bladder disorder
CC that is characterised by inflammation which is associated with symptoms
CC of interstitial cystitis or associated with a disruption of the integrity
CC of the bladder lining. This is the amino acid sequence of a human serum
CC melanocyte stimulating hormone (alpha-MSH) concatamer useful in the
CC treatment of bladder disorders
XX
SQ Sequence 241 AA;

```
  Query Match             100.0%;  Score 123;  DB 7;  Length 241;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48
```

RESULT 16
AAO16985
ID AAO16985 standard; protein; 242 AA.
XX
AC AAO16985;
XX
DT 29-MAY-2002 (first entry)
XX
DE Alpha-MSH construct protein fragment SEQ ID NO: 70.
XX
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX
OS Unidentified.
XX
PN WO200206316-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US022263.
XX
PR 14-JUL-2000; 2000US-0218381P.
PR 18-AUG-2000; 2000US-0226382P.
PR 06-OCT-2000; 2000US-0238380P.
PR 29-DEC-2000; 2000US-0258764P.
PR 14-JUN-2001; 2001US-0298317P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX
DR WPI; 2002-195801/25.
XX
PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX
PS Disclosure; Page 5; 89pp; English.
XX
CC The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention
XX
SQ Sequence 242 AA;

```
  Query Match             100.0%;  Score 123;  DB 5;  Length 242;
  Best Local Similarity   100.0%;  Pred. No. 2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48
```

RESULT 17
AAO16986
ID AAO16986 standard; protein; 244 AA.
XX
AC AAO16986;
XX
DT 29-MAY-2002 (first entry)
XX
DE Alpha-MSH construct protein fragment SEQ ID NO: 71.
XX
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;

KW diabetes; uveitis; coeliac disease.
XX
OS Unidentified.
XX
PN WO200206316-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US022263.
XX
PR 14-JUL-2000; 2000US-0218381P.
PR 18-AUG-2000; 2000US-0226382P.
PR 06-OCT-2000; 2000US-0238380P.
PR 29-DEC-2000; 2000US-0258764P.
PR 14-JUN-2001; 2001US-0298317P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX
DR WPI; 2002-195801/25.
XX
PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX
PS Disclosure; Page 5; 89pp; English.
XX
CC The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention
XX
SQ Sequence 244 AA;

Query Match 100.0%; Score 123; DB 5; Length 244;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 25 DAHKSEVAHRFKDLGEENFKALVL 48

RESULT 18
AAO16987
ID AAO16987 standard; protein; 245 AA.
XX
AC AAO16987;
XX
DT 29-MAY-2002 (first entry)
XX
DE Alpha-MSH construct protein fragment SEQ ID NO: 72.
XX
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX
OS Unidentified.
XX
PN WO200206316-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US022263.
XX
PR 14-JUL-2000; 2000US-0218381P.
PR 18-AUG-2000; 2000US-0226382P.
PR 06-OCT-2000; 2000US-0238380P.
PR 29-DEC-2000; 2000US-0258764P.
PR 14-JUN-2001; 2001US-0298317P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX
DR WPI; 2002-195801/25.
XX
PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX
PS Disclosure; Page 5; 89pp; English.
XX
CC The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention
XX
SQ Sequence 245 AA;

Query Match 100.0%; Score 123; DB 5; Length 245;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 25 DAHKSEVAHRFKDLGEENFKALVL 48

RESULT 19
AAO16988
ID AAO16988 standard; protein; 245 AA.
XX
AC AAO16988;
XX
DT 29-MAY-2002 (first entry)
XX
DE Alpha-MSH construct protein fragment SEQ ID NO: 73.
XX
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX
OS Unidentified.
XX
PN WO200206316-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US022263.
XX
PR 14-JUL-2000; 2000US-0218381P.
PR 18-AUG-2000; 2000US-0226382P.
PR 06-OCT-2000; 2000US-0238380P.
PR 29-DEC-2000; 2000US-0258764P.

PR 14-JUN-2001; 2001US-0298317P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX
DR WPI; 2002-195801/25.
XX
PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX
PS Disclosure; Page 5; 89pp; English.
XX
CC The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention
XX
SQ Sequence 245 AA;

Query Match 100.0%; Score 123; DB 5; Length 245;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 25 DAHKSEVAHRFKDLGEENFKALVL 48

RESULT 20
AAO16989
ID AAO16989 standard; protein; 268 AA.
XX
AC AAO16989;
XX
DT 29-MAY-2002 (first entry)
XX
DE Alpha-MSH construct protein fragment SEQ ID NO: 60.
XX
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX
OS Unidentified.
XX
PN WO200206316-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US022263.
XX
PR 14-JUL-2000; 2000US-0218381P.
PR 18-AUG-2000; 2000US-0226382P.
PR 06-OCT-2000; 2000US-0238380P.
PR 29-DEC-2000; 2000US-0258764P.
PR 14-JUN-2001; 2001US-0298317P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX
DR WPI; 2002-195801/25.
XX
PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX
PS Example 2; Page 5; 89pp; English.
XX
CC The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention
XX
SQ Sequence 268 AA;

Query Match 100.0%; Score 123; DB 5; Length 268;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 25 DAHKSEVAHRFKDLGEENFKALVL 48

RESULT 21
ABU10026
ID ABU10026 standard; protein; 268 AA.
XX
AC ABU10026;
XX
DT 31-JUL-2003 (first entry)
XX
DE Alpha-MSH/Serum albumin fusion protein H4.
XX
KW Bladder disorder; cytostatic; antiinflammatory; immune response;
KW un-methylated CpG sequence; alpha-MSH; melanocortin receptor;
KW bladder cancer; tumour; interstitial cystitis; inflammation;
KW alpha-MSH concatamer; melanocyte stimulating hormone; human;
KW serum albumin; fusion protein; H4.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2002193332-A1.
XX
PD 19-DEC-2002.
XX
PF 12-FEB-2002; 2002US-00074956.
XX
PR 12-FEB-2001; 2001US-0268175P.
XX
PA (HEDL/) HEDLEY M L.
XX
PI Hedley ML;
XX
DR WPI; 2003-447327/42.
XX
PT Modulating immune responses in a mammal with a bladder disorder e.g.
PT bladder cancer, by administering nucleic acids comprising un-methylated
PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to
PT the mammal.
XX
PS Example 2; Page 9; 17pp; English.
XX
CC The invention describes a method of modulating an immune response in a
CC mammal, comprising identifying a mammal that has or is at risk for having
CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)
CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated

CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;
CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The
CC method is useful for modulating immune response in a mammal having a
CC bladder disorder, where administration of (N1) results in an amelioration
CC of one or more symptoms of the disorder. Preferably, the method is useful
CC for modulating immune response in a mammal having bladder cancer (where
CC administration of (N1) results in a decrease in tumour size or activity),
CC or for modulating immune response in a mammal having interstitial
CC cystitis (where administration of (N1) results in a modulation of the
CC immune response from Th2 response to a Th1 response). The method is also
CC useful for modulating immune response in a mammal having bladder disorder
CC that is characterised by inflammation which is associated with symptoms
CC of interstitial cystitis or associated with a disruption of the integrity
CC of the bladder lining. This is the amino acid sequence of a human serum
CC melanocyte stimulating hormone (alpha-MSH) concatamer useful in the
CC treatment of bladder disorders
XX
SQ Sequence 268 AA;

Query Match 100.0%; Score 123; DB 7; Length 268;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 25 DAHKSEVAHRFKDLGEENFKALVL 48

RESULT 22
AAU29581
ID AAU29581 standard; protein; 289 AA.
XX
AC AAU29581;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #72.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 168; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 289 AA;

Query Match 100.0%; Score 123; DB 4; Length 289;
Best Local Similarity 100.0%; Pred. No. 2.4e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 37 DAHKSEVAHRFKDLGEENFKALVL 60

RESULT 23
AAU29575
ID AAU29575 standard; protein; 289 AA.
XX
AC AAU29575;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #66.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 167; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention

```
XX
SQ   Sequence 289 AA;

  Query Match             100.0%;  Score 123;  DB 4;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         37 DAHKSEVAHRFKDLGEENFKALVL 60


RESULT 24
AAR14178
ID   AAR14178 standard; protein; 303 AA.
XX
AC   AAR14178;
XX
DT   19-DEC-1991  (first entry)
XX
DE   Human serum albumin lacking C-terminal fragment.
XX
KW   HSA; fusion protein; drug.
XX
OS   Homo sapiens.
XX
PN   JP03201987-A.
XX
PD   03-SEP-1991.
XX
PF   29-DEC-1989;   89JP-00344701.
XX
PR   29-DEC-1989;   89JP-00344701.
XX
PA   (TOFU ) TONEN CORP.
XX
DR   WPI; 1991-300976/41.
XX
PT   Human serum albumin fragment - where C-terminal of human serum albumin is
PT   lacking and which can be combined with various drugs.
XX
PS   Claim 1; Page 1; 23pp; Japanese.
XX
CC   This sequence corresponds to amino acids 1 to 303 of mature human serum
CC   albumin. The fragment lacking the C-terminal sequence can form part of a
CC   fusion protein, for example with drugs. (This sequence is taken from the
CC   full-length HSA sequence in EP-330451). See also AAR14179
XX
SQ   Sequence 303 AA;

  Query Match             100.0%;  Score 123;  DB 2;  Length 303;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 25
AAP90387
ID   AAP90387 standard; protein; 373 AA.
XX
AC   AAP90387;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   01-NOV-1989  (first entry)
XX
DE   N-terminal of human serum albumin polypeptide.
XX
KW   Human serum albumin polypeptide; plasma expanders.
XX
OS   Homo sapiens; (Human).
XX
PN   EP322094-A.
XX
PD   28-JUN-1989.
XX
PF   25-OCT-1988;   88EP-00310000.
XX
PR   30-OCT-1987;   87GB-00025529.
XX
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
XX
PI   Ballance DJ,  Hinchliffe E,  Geisow MJ,  Senior PJ;
XX
DR   WPI; 1989-186464/26.
XX
PT   New N-terminal fragments of human serum albumin - esp. useful as blood
PT   plasma expanders.
XX
PS   Claim 2; Page 9; 20pp; English.
XX
CC   N-terminal portion of human serum albumin. Used as plasma expanders, or
CC   as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-
CC   MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC   field)
XX
SQ   Sequence 373 AA;

  Query Match             100.0%;  Score 123;  DB 1;  Length 373;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 26
AAP90389
ID   AAP90389 standard; protein; 388 AA.
XX
AC   AAP90389;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   01-NOV-1989  (first entry)
XX
DE   N-terminal human serum albumin polypeptide.
XX
KW   N-terminal human serum albumin polypeptide; plasma expanders.
XX
OS   Homo sapiens; (Human).
XX
PN   EP322094-A.
XX
PD   28-JUN-1989.
XX
PF   25-OCT-1988;   88EP-00310000.
XX
PR   30-OCT-1987;   87GB-00025529.
XX
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
XX
PI   Ballance DJ,  Hinchliffe E,  Geisow MJ,  Senior PJ;
XX
DR   WPI; 1989-186464/26.
XX
PT   New N-terminal fragments of human serum albumin - esp. useful as blood
PT   plasma expanders.
XX
PS   Claim 2; Page 9; 20pp; English.
XX
CC   N-terminal fragment of human serum albumin used as plasma expander, or as
```

```
CC   substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-
CC   2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC   field)
XX
SQ   Sequence 388 AA;

  Query Match             100.0%;  Score 123;  DB 1;  Length 388;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 27
AAP90390
ID   AAP90390 standard; protein; 389 AA.
XX
AC   AAP90390;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   01-NOV-1989  (first entry)
XX
DE   N-terminal human serum albumin.
XX
KW   N-terminal portion of human serum albumin; plasma expanders.
XX
OS   Homo sapiens; (Human).
XX
PN   EP322094-A.
XX
PD   28-JUN-1989.
XX
PF   25-OCT-1988;   88EP-00310000.
XX
PR   30-OCT-1987;   87GB-00025529.
XX
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
XX
PI   Ballance DJ,  Hinchliffe E,  Geisow MJ,  Senior PJ;
XX
DR   WPI; 1989-186464/26.
XX
PT   New N-terminal fragments of human serum albumin - esp. useful as blood
PT   plasma expanders.
XX
PS   Claim 2; Page 9; 20pp; English.
XX
CC   N-terminal portion of human serum albumin. Used to make plasma expanders,
CC   or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25
CC   -MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC   field)
XX
SQ   Sequence 389 AA;

  Query Match             100.0%;  Score 123;  DB 1;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 28
AAP90391
ID   AAP90391 standard; protein; 390 AA.
XX
AC   AAP90391;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   01-NOV-1989  (first entry)
XX
DE   N-terminal human serum albumin.
XX
KW   N-terminal portion of human serum albumin; plasma expanders.
XX
OS   Homo sapiens; (Human).
XX
PN   EP322094-A.
XX
PD   28-JUN-1989.
XX
PF   25-OCT-1988;   88EP-00310000.
XX
PR   30-OCT-1987;   87GB-00025529.
XX
PA   (DELZ ) DELTA BIOTECHNOLOGY LTD.
XX
PI   Ballance DJ,  Hinchliffe E,  Geisow MJ,  Senior PJ;
XX
DR   WPI; 1989-186464/26.
XX
PT   New N-terminal fragments of human serum albumin - esp. useful as blood
PT   plasma expanders.
XX
PS   Claim 2; Page 9; 20pp; English.
XX
CC   N-terminal portion of human serum albumin. Used to make new N-terminal
CC   fragments which are used as plasma expanders, or as substitutes for HSA
CC   or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct PA
CC   field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 390 AA;

  Query Match             100.0%;  Score 123;  DB 1;  Length 390;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 29
AAU29876
ID   AAU29876 standard; protein; 401 AA.
XX
AC   AAU29876;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secreted protein #367.
XX
KW   Human; vaccination; gene therapy; nutritional supplement;
KW   stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW   immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200179449-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US008656.
XX
PR   18-APR-2000; 2000US-00552929.
PR   26-JAN-2001; 2001US-00770160.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
```

DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 206; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 401 AA;

Query Match 100.0%; Score 123; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 3.5e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 38 DAHKSEVAHRFKDLGEENFKALVL 61

RESULT 30
AAP90392
ID AAP90392 standard; protein; 407 AA.
XX
AC AAP90392;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
DE N-terminal human serum albumin.
XX
KW Human serum albumin; mature protein; new polypeptides; plasma expanders.
XX
OS Homo sapiens; (Human).
XX
PN EP322094-A.
XX
PD 28-JUN-1989.
XX
PF 25-OCT-1988; 88EP-00310000.
XX
PR 30-OCT-1987; 87GB-00025529.
XX
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
XX
PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX
DR WPI; 1989-186464/26.
XX
PT New N-terminal fragments of human serum albumin - esp. useful as blood
PT plasma expanders.
XX
PS Claim 2; Page 9; 20pp; English.
XX
CC N-terminal portion of human serum albumin. Used to make plasma expanders,
CC or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25
CC -MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX
SQ Sequence 407 AA;

Query Match 100.0%; Score 123; DB 1; Length 407;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 1 DAHKSEVAHRFKDLGEENFKALVL 24

RESULT 31
ADD32019
ID ADD32019 standard; protein; 500 AA.
XX
AC ADD32019;
XX
DT 15-JAN-2004 (first entry)
XX
DE Heterologous fusion protein related protein sequence SEQ ID NO:25.
XX
KW heterologous fusion protein;
KW hyperglycosylated granulocyte-colony stimulating factor; G-CSF;
KW human albumin; human albumin analogue; immunoglobulin; Fc;
KW immunostimulant; protein therapy; neutrophil level;
KW insufficient circulating neutrophil level;
KW chronic congenital neutropenia.
XX
OS Synthetic.
XX
PN WO2003076567-A2.
XX
PD 18-SEP-2003.
XX
PF 21-FEB-2003; 2003WO-US003120.
XX
PR 05-MAR-2002; 2002US-0361948P.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Beals JM, Kuchibhotla U;
XX
DR WPI; 2003-902770/82.
XX
PT New heterologous fusion proteins with granulocyte-colony stimulating
PT factor activity, useful for increasing neutrophil levels and treating
PT patients with low circulating neutrophils, such as after chemotherapy or
PT in neutropenia.
XX
PS Disclosure; SEQ ID NO 25; 126pp; English.
XX
CC The present invention describes a heterologous fusion protein (I)
CC comprising a hyperglycosylated granulocyte-colony stimulating factor (G-
CC CSF) analogue fused to a polypeptide having human albumin, human albumin
CC analogues, or fragments of human albumin, or the Fc portion of an
CC immunoglobulin, an analogue of the Fc portion of an immunoglobulin, or
CC fragments of the Fc portion of an immunoglobulin. Also described: (1) a
CC heterologous fusion protein, which is the product of the expression in a
CC host cell of an exogenous DNA sequence encoding (I); (2) an isolated
CC nucleic acid sequence comprising: (a) polynucleotides encoding (I); or
CC (b) a polynucleotide, which comprises any of 15 DNA sequences each
CC comprising 1044 base pairs (see ADD31996 to ADD32010), fused to the DNA
CC encoding a protein (i.e. human albumin, human albumin analogue or
CC fragments of human albumin); (3) increasing neutrophil levels in a mammal
CC comprising the administration of (I); (4) pharmaceutical formulations
CC adapted for the treatment of patients with insufficient neutrophil levels
CC comprising any of (I); (5) a vector comprising the polynucleotide of (2);
CC (6) host cells comprising the vector of (5), or expressing at least one
CC heterologous protein; and (7) producing (I). (I) has immunostimulant
CC activity, and can be used in protein therapy. (I) can be used for

CC increasing neutrophil levels in a mammal. (I) are particularly useful in
CC medicaments for the treatment of patients with insufficient circulating
CC neutrophil levels, or for the manufacture of a medicament for the
CC treatment of patients with insufficient circulating neutrophil levels.
CC The fusion protein is particularly useful for treating conditions
CC treatable by stimulation of circulating neutrophils, such as after
CC chemotherapy regimens or in chronic congenital neutropenia. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 500 AA;

Query Match 100.0%; Score 123; DB 7; Length 500;
Best Local Similarity 100.0%; Pred. No. 4.5e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 190 DAHKSEVAHRFKDLGEENFKALVL 213

RESULT 32
ABM84998
ID ABM84998 standard; protein; 544 AA.
XX
AC ABM84998;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:5247.
XX
KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
DR WPI; 2004-329368/30.
DR N-PSDB; ACN43650.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 544 AA;

Query Match 100.0%; Score 123; DB 8; Length 544;
Best Local Similarity 100.0%; Pred. No. 5e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 25 DAHKSEVAHRFKDLGEENFKALVL 48

RESULT 33
AAU29877
ID AAU29877 standard; protein; 550 AA.
XX
AC AAU29877;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #368.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 206; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX

```
SQ   Sequence 550 AA;

  Query Match             100.0%;  Score 123;  DB 4;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         27 DAHKSEVAHRFKDLGEENFKALVL 50


RESULT 34
ABM84997
ID   ABM84997 standard; protein; 571 AA.
XX
AC   ABM84997;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human diagnostic and therapeutic pprotein SEQ ID NO:5246.
XX
KW   gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS   Homo sapiens.
XX
PN   WO2004023973-A2.
XX
PD   25-MAR-2004.
XX
PF   12-SEP-2003; 2003WO-US028227.
XX
PR   12-SEP-2002; 2002US-0410259P.
PR   12-SEP-2002; 2002US-0410260P.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Schmidt JP,  Wright RJ,  Bruns CM,  Marjanovic MM,  Shen F;
PI   Harthshorne TA,  Suchorolski MT,  Altus CM,  Pitts SJ,  Elder LV;
PI   Mooney EM,  Delegeane AM,  Panesar IS,  Banville SC,  Reddy TP;
PI   Stevens KA,  Blanchard JL,  Panzer SR,  Wang X,  Au AP,  Gerstin EH;
PI   Peralta CH,  Anderson SB,  Rioux P,  Shen EJ,  Wu MC,  Stuve LL;
PI   Lagace RE,  Spiro PA,  Stewart EA,  Wingrove J,  Vitt UA,  Kirton ES;
PI   Xu Y,  Kwong M,  Policky JL,  Hurwitz BL,  Ma Y,  Jackson JL,  Gietzen D;
PI   Patury S,  Shi X,  Suarez CJ;
XX
DR   WPI; 2004-329368/30.
DR   N-PSDB; ACN43649.
XX
PT   New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT   in diagnosing a condition, disease or disorder associated with human
PT   molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT   in gene mapping.
XX
PS   Claim 27; Page; 190pp; English.
XX
CC   The invention relates to novel diagnostic and therapeutic polynucleotides
CC   selected from one of the 2722 sequences defined in the specification. A
CC   polynucleotide of the invention may have a use in gene therapy. The human
CC   diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC   used to diagnose a particular condition, disease or disorder associated
CC   with human molecules, e.g. cell proliferative disorders,
CC   autoimmune/inflammatory disorder, developmental disorder, endocrine
CC   disorder, neurological disorders, gastrointestinal disorders, or
CC   infections caused by virus, bacteria, fungi or parasite. The dithp
CC   molecules may also be used in genetic mapping, in identifying individuals
CC   from minute biological samples, in detecting single nucleotide
CC   polymorphisms, as molecular weight markers, and for somatic or germline
CC   gene therapy. The present sequence represents a dithp protein of the
CC   invention. Note: The sequence data for this patent is not represented in
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ   Sequence 571 AA;

  Query Match             100.0%;  Score 123;  DB 8;  Length 571;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 35
ABM84996
ID   ABM84996 standard; protein; 573 AA.
XX
AC   ABM84996;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human diagnostic and therapeutic pprotein SEQ ID NO:5245.
XX
KW   gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS   Homo sapiens.
XX
PN   WO2004023973-A2.
XX
PD   25-MAR-2004.
XX
PF   12-SEP-2003; 2003WO-US028227.
XX
PR   12-SEP-2002; 2002US-0410259P.
PR   12-SEP-2002; 2002US-0410260P.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Schmidt JP,  Wright RJ,  Bruns CM,  Marjanovic MM,  Shen F;
PI   Harthshorne TA,  Suchorolski MT,  Altus CM,  Pitts SJ,  Elder LV;
PI   Mooney EM,  Delegeane AM,  Panesar IS,  Banville SC,  Reddy TP;
PI   Stevens KA,  Blanchard JL,  Panzer SR,  Wang X,  Au AP,  Gerstin EH;
PI   Peralta CH,  Anderson SB,  Rioux P,  Shen EJ,  Wu MC,  Stuve LL;
PI   Lagace RE,  Spiro PA,  Stewart EA,  Wingrove J,  Vitt UA,  Kirton ES;
PI   Xu Y,  Kwong M,  Policky JL,  Hurwitz BL,  Ma Y,  Jackson JL,  Gietzen D;
PI   Patury S,  Shi X,  Suarez CJ;
XX
DR   WPI; 2004-329368/30.
DR   N-PSDB; ACN43648.
XX
PT   New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT   in diagnosing a condition, disease or disorder associated with human
PT   molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT   in gene mapping.
XX
PS   Claim 27; Page; 190pp; English.
XX
CC   The invention relates to novel diagnostic and therapeutic polynucleotides
CC   selected from one of the 2722 sequences defined in the specification. A
CC   polynucleotide of the invention may have a use in gene therapy. The human
CC   diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC   used to diagnose a particular condition, disease or disorder associated
CC   with human molecules, e.g. cell proliferative disorders,
CC   autoimmune/inflammatory disorder, developmental disorder, endocrine
CC   disorder, neurological disorders, gastrointestinal disorders, or
CC   infections caused by virus, bacteria, fungi or parasite. The dithp
CC   molecules may also be used in genetic mapping, in identifying individuals
CC   from minute biological samples, in detecting single nucleotide
CC   polymorphisms, as molecular weight markers, and for somatic or germline
CC   gene therapy. The present sequence represents a dithp protein of the
CC   invention. Note: The sequence data for this patent is not represented in
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ   Sequence 573 AA;
```

```
  Query Match             100.0%;  Score 123;  DB 8;  Length 573;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 36
ABM84995
ID   ABM84995 standard; protein; 573 AA.
XX
AC   ABM84995;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human diagnostic and therapeutic pprotein SEQ ID NO:5244.
XX
KW   gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS   Homo sapiens.
XX
PN   WO2004023973-A2.
XX
PD   25-MAR-2004.
XX
PF   12-SEP-2003; 2003WO-US028227.
XX
PR   12-SEP-2002; 2002US-0410259P.
PR   12-SEP-2002; 2002US-0410260P.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Schmidt JP,  Wright RJ,  Bruns CM,  Marjanovic MM,  Shen F;
PI   Harthshorne TA,  Suchorolski MT,  Altus CM,  Pitts SJ,  Elder LV;
PI   Mooney EM,  Delegeane AM,  Panesar IS,  Banville SC,  Reddy TP;
PI   Stevens KA,  Blanchard JL,  Panzer SR,  Wang X,  Au AP,  Gerstin EH;
PI   Peralta CH,  Anderson SB,  Rioux P,  Shen EJ,  Wu MC,  Stuve LL;
PI   Lagace RE,  Spiro PA,  Stewart EA,  Wingrove J,  Vitt UA,  Kirton ES;
PI   Xu Y,  Kwong M,  Policky JL,  Hurwitz BL,  Ma Y,  Jackson JL,  Gietzen D;
PI   Patury S,  Shi X,  Suarez CJ;
XX
DR   WPI; 2004-329368/30.
DR   N-PSDB; ACN43647.
XX
PT   New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT   in diagnosing a condition, disease or disorder associated with human
PT   molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT   in gene mapping.
XX
PS   Claim 27; Page; 190pp; English.
XX
CC   The invention relates to novel diagnostic and therapeutic polynucleotides
CC   selected from one of the 2722 sequences defined in the specification. A
CC   polynucleotide of the invention may have a use in gene therapy. The human
CC   diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC   used to diagnose a particular condition, disease or disorder associated
CC   with human molecules, e.g. cell proliferative disorders,
CC   autoimmune/inflammatory disorder, developmental disorder, endocrine
CC   disorder, neurological disorders, gastrointestinal disorders, or
CC   infections caused by virus, bacteria, fungi or parasite. The dithp
CC   molecules may also be used in genetic mapping, in identifying individuals
CC   from minute biological samples, in detecting single nucleotide
CC   polymorphisms, as molecular weight markers, and for somatic or germline
CC   gene therapy. The present sequence represents a dithp protein of the
CC   invention. Note: The sequence data for this patent is not represented in
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ   Sequence 573 AA;

  Query Match             100.0%;  Score 123;  DB 8;  Length 573;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 37
ABG72381
ID   ABG72381 standard; protein; 584 AA.
XX
AC   ABG72381;
XX
DT   10-FEB-2003  (first entry)
XX
DE   Mature human serum albumin #2.
XX
KW   Human; serum albumin; HSA; cancer; cytostatic; breast cancer;
KW   prostate cancer; anti-proliferative.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 241. .242
FT                   /note= "Encoded by GTCCACACG"
XX
PN   US2002123080-A1.
XX
PD   05-SEP-2002.
XX
PF   14-AUG-2001; 2001US-00929552.
XX
PR   19-DEC-1996;   96US-00769746.
XX
PA   (TUFT ) UNIV TUFTS.
XX
PI   Sonnenschein C,  Soto AM;
XX
DR   WPI; 2003-066789/06.
DR   N-PSDB; ABX13582.
XX
PT   Testing human cancer cells, especially breast and prostate cancer cells,
PT   by contacting cancer cells obtained from biopsy of a patient ex vivo with
PT   antibody specific to human albumin receptor.
XX
PS   Disclosure; Fig 2; 17pp; English.
XX
CC   The invention relates to testing human cancer cells, comprising obtaining
CC   cancer cells from the patient and contacting the cell ex vivo with an
CC   antibody to the receptor for human albumin. The method is useful for
CC   testing human cancer cells in particular breast and prostate cancer
CC   cells. The present sequence is mature human serum albumin, HSA. The anti-
CC   proliferative effect of HSA was assayed in an experiment included in the
CC   specification
XX
SQ   Sequence 584 AA;

  Query Match             100.0%;  Score 123;  DB 6;  Length 584;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKALVL 24


RESULT 38
AAP93344
ID   AAP93344 standard; protein; 585 AA.
XX
AC   AAP93344;
XX
```

DT 25-MAR-2003 (revised)
DT 23-JUN-1990 (first entry)
XX
DE Sequence of mature human serum albumin (HSA) as encoded by artificial
DE gene.
XX
KW Mature human serum albumin; artificial gene; oligonucleotide block;
KW hypobolaemia; shock; hypoalbuminaemia.
XX
OS Homo sapiens.
XX
PN EP308381-A.
XX
PD 22-MAR-1989.
XX
PF 13-SEP-1988; 88EP-00850299.
XX
PR 14-SEP-1987; 87SE-00003539.
XX
PA (SKAN-) SKANDIGEN AB.
PA (MAGY) MTA SZEGEDI BIOLOG KOEZPONTI.
PA (VEPE-) VEPEX CONTRACTOR LTD.
XX
PI Aberg B, Simoncsits A, Kalan M, Csperpan I, Bajszar G;
XX
DR WPI; 1989-087749/12.
DR N-PSDB; AAN90997.
XX
PT Artificial gene coding for authentic human serum albumin - constructed on
PT the basis of codons most frequently used by chosen non-human host.
XX
PS Disclosure; pp. 11-16; 121pp; English.
XX
CC The synthetic gene was constructed by designing a nucleotide sequence in
CC which the codons which are most frequently used by the chosen non-human
CC host were selected. In this case, it is yeast cells (LL20; Leu2-3; 112,
CC His 3-11, 15). The synthetic HSA gene was assembled from 24
CC oligonucleotide blocks. HSA is used in therapy for the treatment of
CC hypovolaemia, shock and hypoalbuminaemia. It is also used as an additive
CC in perfusion liq. for extracorporeal circulation and as an experimental
CC antigen. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR
CC -2003 to correct PI field.)
XX
SQ Sequence 585 AA;

Query Match 100.0%; Score 123; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 1 DAHKSEVAHRFKDLGEENFKALVL 24

RESULT 39
AAP90388
ID AAP90388 standard; protein; 585 AA.
XX
AC AAP90388;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Mature human serum albumin polypeptide.
XX
KW Human serum albumin; mature protein; new polypeptides; plasma expanders.
XX
OS Homo sapiens; (Human).
XX
PN EP322094-A.
XX
PD 28-JUN-1989.
XX
PF 25-OCT-1988; 88EP-00310000.
XX
PR 30-OCT-1987; 87GB-00025529.
XX
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
XX
PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX
DR WPI; 1989-186464/26.
DR N-PSDB; AAN90128.
XX
PT New N-terminal fragments of human serum albumin - esp. useful as blood
PT plasma expanders.
XX
PS Disclosure; Fig 2; 20pp; English.
XX
CC Mature protein of human serum albumin (see corresp. AAN90128). Used to
CC make new N-terminal fragments which are used as plasma expanders, or as
CC substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX
SQ Sequence 585 AA;

Query Match 100.0%; Score 123; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 1 DAHKSEVAHRFKDLGEENFKALVL 24

RESULT 40
AAP91422
ID AAP91422 standard; protein; 585 AA.
XX
AC AAP91422;
XX
DT 25-MAR-2003 (revised)
DT 20-DEC-1989 (first entry)
XX
DE Human normal serum albumin A.
XX
KW Human normal serum albumin A; pAT-pho-HSA-A; haemorrhagic shock;
KW hypoalbuminaemia.
XX
OS Homo sapiens.
XX
PN EP330451-A.
XX
PD 30-AUG-1989.
XX
PF 22-FEB-1989; 89EP-00301731.
XX
PR 22-FEB-1988; 88JP-00037453.
XX
PA (TOFU) TOA NENRYO KOGYO KK.
XX
PI Suzuki M, Maki N, Yagi S;
XX
DR WPI; 1989-250534/35.
XX
PT CDNA encoding human normal serum albumin contained in plasmid - obtd. by
PT culturing host transformed with expression vector comprising CDNA coding
PT for albumin.
XX
PS Claim 1; Fig 3-1 to 3-5; 19pp; English.
XX
CC cDNA amino acid sequence of human serum albumin A (HSA-A) which is
CC identical to that encoded by chromosomal DNA. Previous polypeptides
CC produced from cDNA have one or more amino acids which differ from those

CC of HSA-A produced from the chromosomal DNA, and may exhibit antigenicity
CC when administered to humans. The HSA-A is used to treat haemorrhagic
CC shock and hypoalbuminaemia. See also AAN90600. (Updated on 25-MAR-2003 to
CC correct PI field.)
XX
SQ Sequence 585 AA;

Query Match 100.0%; Score 123; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 1 DAHKSEVAHRFKDLGEENFKALVL 24

Search completed: August 19, 2005, 10:58:24
Job time : 165 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 10:45:00 ; Search time 56 Seconds
(without alignments)
219.463 Million cell updates/sec

Title: US-09-846-328B-1_COPY_2_25
Perfect score: 123
Sequence: 1 DAHKSEVAHRFKDLGEENFKALVL 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	123	100.0	167	2	Q6UXK4	Q6uxk4 homo sapien
2	123	100.0	396	2	Q8IUK7	Q8iuk7 homo sapien
3	123	100.0	609	1	ALBU_HUMAN	P02768 homo sapien
4	117	95.1	609	2	Q68DN5	Q68dn5 homo sapien
5	111	90.2	21	2	Q9QVA1	Q9qva1 rattus sp.
6	111	90.2	417	2	Q86YG0	Q86yg0 homo sapien
7	110	89.4	600	1	ALBU_MACMU	Q28522 macaca mula
8	110	89.4	608	2	Q95VB7	Q95vb7 schistosoma
9	109	88.6	607	1	ALBU_BOVIN	P02769 bos taurus
10	106	86.2	608	1	ALBU_RAT	P02770 rattus norv
11	105	85.4	607	1	ALBU_SHEEP	P14639 ovis aries
12	100	81.3	607	1	ALBU_HORSE	P35747 equus cabal
13	98	79.7	205	2	Q8CG74	Q8cg74 mus musculu
14	98	79.7	608	1	ALBU_MOUSE	P07724 mus musculu
15	98	79.7	608	2	Q8C7H3	Q8c7h3 mus musculu
16	97	78.9	584	2	Q7YSG3	Q7ysg3 felis silve
17	97	78.9	605	1	ALBU_PIG	P08835 sus scrofa
18	97	78.9	607	2	Q68NH7	Q68nh7 sus scrofa
19	97	78.9	608	1	ALBU_FELCA	P49064 felis silve
20	97	78.9	608	2	Q6WDN9	Q6wdn9 cavia porce
21	97	78.9	609	1	ALBU_MERUN	O35090 meriones un
22	93	75.6	40	2	Q9TRA5	Q9tra5 oryctolagus
23	93	75.6	608	1	ALBU_RABIT	P49065 oryctolagus
24	92	74.8	20	2	Q9QUX8	Q9qux8 rattus sp.
25	92	74.8	608	1	ALBU_CANFA	P49822 canis famil
26	89	72.4	30	2	Q7M3A0	Q7m3a0 tachyglossu
27	87	70.7	20	2	Q9R4X7	Q9r4x7 mycobacteri
28	87	70.7	25	2	Q9TRW8	Q9trw8 trichosurus
29	84	68.3	583	2	Q6B3Z0	Q6b3z0 elephas max
30	81	65.9	615	1	ALBU_CHICK	P19121 gallus gall
31	76	61.8	40	1	ALB1_TRASC	P81188 trachemys s
32	71	57.7	17	2	Q9TR98	Q9tr98 canis famil
33	71	57.7	30	2	Q9PRW0	Q9prw0 struthio ca
34	61	49.6	144	2	Q8E1L1	Q8e1l1 streptococc
35	61	49.6	144	2	Q8E728	Q8e728 streptococc
36	60	48.8	15	1	NF41_NAEFO	P83729 naegleria f
37	59	48.0	548	2	Q7M8S9	Q7m8s9 wolinella s
38	58	47.2	101	1	ALBU_NEOFS	P83517 neoceratodu
39	56	45.5	576	2	Q8C7C7	Q8c7c7 mus musculu
40	53	43.1	569	2	Q6ZQ53	Q6zq53 mus musculu
41	53	43.1	992	2	Q7TT14	Q7tt14 mus musculu
42	53	43.1	1002	1	RBMC_MOUSE	Q8r4x3 mus musculu
43	53	43.1	1032	2	Q6XLI7	Q6xli7 rattus norv
44	52	42.3	112	2	Q7NJ13	Q7nj13 gloeobacter
45	52	42.3	384	2	Q83K42	Q83k42 shigella fl

ALIGNMENTS

RESULT 1
Q6UXK4
ID Q6UXK4 PRELIMINARY; PRT; 167 AA.
AC Q6UXK4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ALB (Fragment).
GN ORFNames=UNQ696;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
DR EMBL; AY358313; AAQ89947.1; -.
DR HSSP; P02768; 1AO6.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; Serum_albumin; 1.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 1.
FT NON_TER 167 167
SQ SEQUENCE 167 AA; 19344 MW; 70AAC3DDA754491C CRC64;

Query Match 100.0%; Score 123; DB 2; Length 167;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48
```

RESULT 2
Q8IUK7
ID Q8IUK7 PRELIMINARY; PRT; 396 AA.
AC Q8IUK7;

```
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   ALB protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC035969; AAH35969.1; -.
DR   HSSP; P02768; 1N5U.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 2.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   SMART; SM00103; ALBUMIN; 2.
DR   PROSITE; PS00212; ALBUMIN; 2.
SQ   SEQUENCE   396 AA;  45159 MW;  756519C096463A9B CRC64;

  Query Match           100.0%;  Score 123;  DB 2;  Length 396;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-09;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 3
ALBU_HUMAN
ID   ALBU_HUMAN     STANDARD;      PRT;   609 AA.
AC   P02768; O95574; Q13140; Q9P157; Q9P1I7; Q9UHS3; Q9UJZ0;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Serum albumin precursor.
GN   Name=ALB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86196112; PubMed=3009475;
RA   Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA   Beattie W.G., Dugaiczyk A.;
RT   "Molecular structure of the human albumin gene is revealed by
RT   nucleotide sequence within q11-22 of chromosome 4.";
RL   J. Biol. Chem. 261:6747-6757(1986).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT LYS-420.
RX   MEDLINE=82081882; PubMed=6171778;
RA   Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA   Najarian R.C., Seeburg P.H., Wion K.L.;
RT   "The sequence of human serum albumin cDNA and its expression in E.
RT   coli.";
RL   Nucleic Acids Res. 9:6103-6114(1981).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANT GLY-121.
RX   MEDLINE=82105994; PubMed=6275391;
RA   Dugaiczyk A., Law S.W., Dennison O.E.;
RT   "Nucleotide sequence and the encoded amino acids of human serum
RT   albumin mRNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).
RC   TISSUE=Fetal liver;
RA   Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
RA   Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
RT   "Functional prediction of the coding sequences of 121 new genes
RT   deduced by analysis of cDNA clones from human fetal liver.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
RA   Huang M.C., Wu H.T.;
RT   "The cDNA sequences of human serum albumin.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver, and Skeletal muscle;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   SEQUENCE OF 25-609.
RX   MEDLINE=76187907; PubMed=1225573; DOI=10.1016/0014-5793(75)80242-0;
RA   Meloun B., Moravek L., Kostka V.;
RT   "Complete amino acid sequence of human serum albumin.";
RL   FEBS Lett. 58:134-137(1975).
RN   [9]
RP   SEQUENCE OF 25-609.
RA   Brown J.R., Shockley P., Behrens P.Q.;
RL   (In) Bing D.H. (eds.);
```

```
RL   The chemistry and physiology of the human plasma proteins, pp.23-40,
RL   Pergamon Press, New York (1979).
RN   [10]
RP   SEQUENCE OF 1-455 FROM N.A.
RC   TISSUE=Liver;
RA   Menaya J., Parrilla R., Ayuso M.S.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
RN   [11]
RP   SEQUENCE OF 1-26 FROM N.A.
RX   MEDLINE=86140099; PubMed=2419329;
RA   Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT   "The human albumin gene. Characterization of the 5' and 3' flanking
RT   regions and the polymorphic gene transcripts.";
RL   J. Biol. Chem. 261:3244-3251(1986).
RN   [12]
RP   SEQUENCE OF 222-229.
RX   MEDLINE=76257808; PubMed=955075; DOI=10.1016/0014-5793(76)80496-6;
RA   Walker J.E.;
RT   "Lysine residue 199 of human serum albumin is modified by
RT   acetylsalicylic acid.";
RL   FEBS Lett. 66:173-175(1976).
RN   [13]
RP   SEQUENCE OF 25-44 AND 480-499.
RC   TISSUE=Heart;
RX   MEDLINE=95203287; PubMed=7895732;
RA   Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT   "The human myocardial two-dimensional gel protein database: update
RT   1994.";
RL   Electrophoresis 15:1459-1465(1994).
RN   [14]
RP   DISULFIDE BONDS.
RA   Saber M.A., Stockbauer P., Moravek L., Meloun B.;
RT   "Disulfide bonds in human serum albumin.";
RL   Collect. Czech. Chem. Commun. 42:564-579(1977).
RN   [15]
RP   BILIRUBIN-BINDING SITE.
RX   MEDLINE=78186630; PubMed=656055;
RA   Jacobsen C.;
RT   "Lysine residue 240 of human serum albumin is involved in high-
RT   affinity binding of bilirubin.";
RL   Biochem. J. 171:453-459(1978).
RN   [16]
RP   VARIANT CANTERBURY ASN-337.
RX   MEDLINE=87157744; PubMed=3828358; DOI=10.1016/0167-4838(87)90088-4;
RA   Brennan S.O., Herbert P.;
RT   "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
RT   domain of serum albumin.";
RL   Biochim. Biophys. Acta 912:191-197(1987).
RN   [17]
RP   VARIANTS NAG-2 AND NAG-3.
RX   MEDLINE=88068523; PubMed=3479777;
RA   Takshashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
RA   Satoh C., Neel J.V.;
RT   "Amino acid substitutions in inherited albumin variants from
RT   Amerindian and Japanese populations.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN   [18]
RP   VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RX   MEDLINE=89345611; PubMed=2762316;
RA   Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,
RA   Neel J.V., Sakurabayashi I., Putnam F.W.;
RT   "Point substitutions in Japanese alloalbumins.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN   [19]
RP   VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RX   MEDLINE=90115905; PubMed=2404284;
RA   Arai K., Madison J., Shimuzu A., Putnam F.W.;
RT   "Point substitutions in albumin genetic variants from Asia.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN   [20]
RP   DESCRIPTION OF VARIANT REDHILL.
RX   MEDLINE=90115852; PubMed=2104980;
RA   Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RT   "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
RT   human serum albumin whose precursor has an aberrant signal peptidase
RT   cleavage site.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN   [21]
RP   VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX   MEDLINE=91062352; PubMed=2247440;
RA   Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA   Watkins S., Putnam F.W.;
RT   "Mutations in genetic variants of human serum albumin found in
RT   Italy.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN   [22]
RP   VARIANT VENEZIA.
RX   MEDLINE=91296740; PubMed=2068071;
RA   Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA   Minchiotti L., Putnam F.W.;
RT   "A donor splice mutation and a single-base deletion produce two
RT   carboxyl-terminal variants of human serum albumin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN   [23]
RP   VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RP   KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
RX   MEDLINE=92052189; PubMed=1946412;
RA   Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA   Matsuda Y.-I., Amaki I., Putnam F.W.;
RT   "Genetic variants of serum albumin in Americans and Japanese.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
RN   [24]
RP   VARIANT CASEBROOK ASN-518.
RX   MEDLINE=91316157; PubMed=1859851; DOI=10.1016/0925-4439(91)90023-3;
RA   Peach R.J., Brennan S.O.;
RT   "Structural characterization of a glycoprotein variant of human serum
RT   albumin: albumin Casebrook (494 Asp-->Asn).";
RL   Biochim. Biophys. Acta 1097:49-54(1991).
RN   [25]
RP   VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX   MEDLINE=92190239; PubMed=1347703; DOI=10.1016/0167-4838(92)90207-T;
RA   Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA   Rochu D., Porta F.;
RT   "Two alloalbumins with identical electrophoretic mobility are produced

  Query Match             100.0%;  Score 123;  DB 1;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-09;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKALVL 48


RESULT 4
Q68DN5
ID   Q68DN5      PRELIMINARY;      PRT;   609 AA.
AC   Q68DN5;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein DKFZp779N1935.
GN   Name=DKFZp779N1935;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RG   The German cDNA Consortium;
RA   Bloecker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA   Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR749331; CAH18185.1; -.
DR   InterPro; IPR001703; Alphafetoprot.
```

```
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Hypothetical protein.
SQ   SEQUENCE   609 AA;  69402 MW;  3BA3AFF17BF99E94 CRC64;

  Query Match             95.1%;  Score 117;  DB 2;  Length 609;
  Best Local Similarity   95.8%;  Pred. No. 1.3e-08;
  Matches   23;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||:||||||||||||||||||||||
Db         25 DAYKSEVAHRFKDLGEENFKALVL 48


RESULT 5
Q9QVA1
ID   Q9QVA1      PRELIMINARY;      PRT;    21 AA.
AC   Q9QVA1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   49 kDa protein (Fragment).
OS   Rattus sp. .
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93105636; PubMed=1468220;
RA   Zhu L., Crouch R.K.;
RT   "Albumin in the cornea is oxidized by hydrogen peroxide.";
RL   Cornea 11:567-572(1992).
DR   HSSP; P02768; 1N5U.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 1.
SQ   SEQUENCE   21 AA;  2429 MW;  26134A3D7CE29FAC CRC64;

  Query Match             90.2%;  Score 111;  DB 2;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-09;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKA 21
              |||||||||||||||||||||
Db          1 DAHKSEVAHRFKDLGEENFKA 21


RESULT 6
Q86YG0
ID   Q86YG0      PRELIMINARY;      PRT;   417 AA.
AC   Q86YG0;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   ALB protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC041789; AAH41789.1; -.
DR   HSSP; P02768; 1E7B.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 2.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 2.
DR   PROSITE; PS00212; ALBUMIN; 2.
SQ   SEQUENCE   417 AA;  47360 MW;  16E764833EEF4E8D CRC64;

  Query Match             90.2%;  Score 111;  DB 2;  Length 417;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-08;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKA 21
              |||||||||||||||||||||
Db         25 DAHKSEVAHRFKDLGEENFKA 45


RESULT 7
ALBU_MACMU
ID   ALBU_MACMU      STANDARD;      PRT;   600 AA.
AC   Q28522;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor (Fragment).
GN   Name=ALB;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93211971; PubMed=8460152;
RA   Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA   Dwulet J., Putnam F.W.;
RT   "cDNA and protein sequence of polymorphic macaque albumins that differ
RT   in bilirubin binding.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
```

```
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M90463; AAA36906.1; -.
DR   PIR; A47391; A47391.
DR   HSSP; P02768; 1E7B.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     10       By similarity.
FT   PROPEP       11     16       By similarity.
FT   CHAIN        17    600       Serum albumin.
FT   DOMAIN       17    197       Albumin 1.
FT   DOMAIN      204    389       Albumin 2.
FT   DOMAIN      396    587       Albumin 3.
FT   METAL        19     19       Copper (By similarity).
FT   BINDING     256    256       Bilirubin (Potential).
FT   DISULFID     69     78       By similarity.
FT   DISULFID     91    107       By similarity.
FT   DISULFID    106    117       By similarity.
FT   DISULFID    140    185       By similarity.
FT   DISULFID    184    193       By similarity.
FT   DISULFID    216    262       By similarity.
FT   DISULFID    261    269       By similarity.
FT   DISULFID    281    295       By similarity.
FT   DISULFID    294    305       By similarity.
FT   DISULFID    332    377       By similarity.
FT   DISULFID    376    385       By similarity.
FT   DISULFID    408    454       By similarity.
FT   DISULFID    453    464       By similarity.
FT   DISULFID    477    493       By similarity.
FT   DISULFID    492    503       By similarity.
FT   DISULFID    530    575       By similarity.
FT   DISULFID    574    583       By similarity.
SQ   SEQUENCE   600 AA;  67880 MW;  E45C871A670E740B CRC64;

  Query Match             89.4%;  Score 110;  DB 1;  Length 600;
  Best Local Similarity   87.5%;  Pred. No. 1.4e-07;
  Matches   21;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||||||||||||||:|| |||
Db         17 DTHKSEVAHRFKDLGEEHFKGLVL 40


RESULT 8
Q95VB7
ID   Q95VB7      PRELIMINARY;      PRT;   608 AA.
AC   Q95VB7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Albumin.
OS   Schistosoma mansoni (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
OX   NCBI_TaxID=6183;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF418550; AAL08579.1; -.
DR   HSSP; P02768; 1HK1.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 2.
SQ   SEQUENCE   608 AA;  68225 MW;  E5EABB28E1C66E54 CRC64;

  Query Match             89.4%;  Score 110;  DB 2;  Length 608;
  Best Local Similarity   83.3%;  Pred. No. 1.4e-07;
  Matches   20;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              ||||||:|||||||||::|| |||
Db         25 DAHKSEIAHRFKDLGEQHFKGLVL 48


RESULT 9
ALBU_BOVIN
ID   ALBU_BOVIN     STANDARD;      PRT;   607 AA.
AC   P02769; O02787;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Serum albumin precursor (Allergen Bos d 6) (BSA).
GN   Name=ALB;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Holowachuk E.W., Stoltenborg J.K., Reed R.G., Peters T. Jr.;
RT   "Bovine serum albumin: cDNA sequence and expression.";
RL   Submitted (AUG-1991) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT THR-214.
RC   TISSUE=Liver;
RA   Barry T., Power S., Gannon F.;
RT   "The bovine serum albumin mRNA.";
RL   Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=21195144; PubMed=11298124;
RA   Hilger C., Grigioni F., De Beaufort C., Michel G., Freilinger J.,
RA   Hentges F.;
RT   "Differential binding of IgG and IgA antibodies to antigenic
RT   determinants of bovine serum albumin.";
RL   Clin. Exp. Immunol. 123:387-394(2001).
RN   [4]
RP   SEQUENCE FROM N.A., AND VARIANT THR-214.
RA   Wu H.T., Huang M.C.;
RT   "The complete cDNA sequence of bovine serum albumin.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 1-32.
RX   MEDLINE=80024278; PubMed=488109;
RA   McGillivray R.T.A., Chung D.W., Davie E.W.;
RT   "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
RT   terminal sequence of preproalbumin.";
RL   Eur. J. Biochem. 98:477-485(1979).
RN   [6]
RP   SEQUENCE OF 19-28.
RX   MEDLINE=77134075; PubMed=843354;
```

```
RA   Patterson J.E., Geller D.M.;
RT   "Bovine microsomal albumin: amino terminal sequence of bovine
RT   proalbumin.";
RL   Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
RN   [7]
RP   SEQUENCE, AND REVISIONS TO 118-119 AND 180.
RX   MEDLINE=91083649; PubMed=2260975;
RA   Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;
RT   "Rapid confirmation and revision of the primary structure of bovine
RT   serum albumin by ESIMS and Frit-FAB LC/MS.";
RL   Biochem. Biophys. Res. Commun. 173:639-646(1990).
RN   [8]
RP   SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
RA   Brown J.R.;
RT   "Structure of bovine serum albumin.";
RL   Fed. Proc. 34:591-591(1975).
RN   [9]
RP   REVISIONS TO 190-195.
RA   Brown J.R.;
RL   Submitted (APR-1975) to the PIR data bank.
RN   [10]
RP   SEQUENCE OF 25-64.
RX   PubMed=2379503;
RA   Strawich E., Glimcher M.J.;
RT   "Tooth 'enamelins' identified mainly as serum proteins. Major
RT   'enamelin' is albumin.";
RL   Eur. J. Biochem. 191:47-56(1990).
RN   [11]
RP   SEQUENCE OF 25-41.
RX   MEDLINE=88267456; PubMed=3389500;
RA   Hsieh J.C., Lin F.P., Tam M.F.;
RT   "Electroblotting onto glass-fiber filter from an analytical
RT   isoelectrofocusing gel: a preparative method for isolating proteins
RT   for N-terminal microsequencing.";
RL   Anal. Biochem. 170:1-8(1988).
RN   [12]
RP   SEQUENCE OF 163-172.
RX   PubMed=2474609;
RA   Carraway R.E., Cochrane D.E., Boucher W., Mitra S.P.;
RT   "Structures of histamine-releasing peptides formed by the action of
RT   acid proteases on mammalian albumin(s).";
RL   J. Immunol. 143:1680-1684(1989).
RN   [13]
RP   SEQUENCE OF 402-433.
RX   MEDLINE=82023364; PubMed=7283978;
RA   Reed R.G., Putnam F.W., Peters T. Jr.;
RT   "Sequence of residues 400-403 of bovine serum albumin.";
RL   Biochem. J. 191:867-868(1980).
RN   [14]
RP   SEQUENCE OF 437-451.
RA   Vilbois F.;
RL   Submitted (AUG-1998) to Swiss-Prot.
RN   [15]
RP   DISULFIDE BONDS.
RA   Brown J.R.;
RT   "Structure of serum albumin: disulfide bridges.";
RL   Fed. Proc. 33:1389-1389(1974).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- ALLERGEN: Causes an allergic reaction in human.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M73993; AAA51411.1; -.
DR   EMBL; X58989; CAA41735.1; -.
DR   EMBL; Y17769; CAA76847.1; -.
DR   EMBL; AF542068; AAN17824.1; -.
DR   PIR; A38885; ABBOS.
DR   HSSP; P02768; 1HK1.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Allergen; Copper; Direct protein sequencing; Lipid-binding;
KW   Metal-binding; Polymorphism; Repeat; Signal.
FT   SIGNAL        1     18
FT   PROPEP       19     24
FT   CHAIN        25    607       Serum albumin.
FT   DOMAIN       25    204       Albumin 1.
FT   DOMAIN      211    396       Albumin 2.
FT   DOMAIN      403    594       Albumin 3.
FT   METAL        27     27       Copper (By similarity).
FT   DISULFID     77     86
FT   DISULFID     99    115
FT   DISULFID    114    125
FT   DISULFID    147    192
FT   DISULFID    191    200
FT   DISULFID    223    269
FT   DISULFID    268    276
FT   DISULFID    288    302
FT   DISULFID    301    312
FT   DISULFID    339    384
FT   DISULFID    383    392
FT   DISULFID    415    461
FT   DISULFID    460    471
FT   DISULFID    484    500
FT   DISULFID    499    510
FT   DISULFID    537    582
FT   DISULFID    581    590
FT   VARIANT     214    214       A -> T.
FT   CONFLICT     58     58       Missing (in Ref. 10).
FT   CONFLICT    302    302       C -> K (in Ref. 8).
FT   CONFLICT    304    305       KP -> PC (in Ref. 8).
FT   CONFLICT    324    324       N -> D (in Ref. 8).
FT   CONFLICT    394    395       ST -> TS (in Ref. 8).
FT   CONFLICT    437    437       K -> R (in Ref. 14).
FT   CONFLICT    493    494       SE -> ES (in Ref. 8).
SQ   SEQUENCE   607 AA;  69293 MW;  39167DFE768585D4 CRC64;

  Query Match             88.6%;  Score 109;  DB 1;  Length 607;
  Best Local Similarity   83.3%;  Pred. No. 2e-07;
  Matches   20;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||:|||||||||||:|| |||
Db         25 DTHKSEIAHRFKDLGEEHFKGLVL 48


RESULT 10
ALBU_RAT
ID   ALBU_RAT       STANDARD;      PRT;   608 AA.
AC   P02770; P11382;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN   Name=Alb;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81223722; PubMed=7017712;
RA   Sargent T.D., Yang M., Bonner J.;
RT   "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
RN   [2]
RP   SEQUENCE OF 1-38, AND PROCESSING.
RX   MEDLINE=77249657; PubMed=893447;
RA   Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
RT   "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
RT   piece. Analysis of the direct translation product of albumin messenger
RT   RNA.";
RL   J. Biol. Chem. 252:6846-6855(1977).
RN   [3]
RP   SEQUENCE OF 25-222.
RX   MEDLINE=78109429; PubMed=564345;
RA   Isemura S., Ikenaka T.;
RT   "Amino acid sequences of fragments I and II obtained by cyanogen
RT   bromide cleavage of rat serum albumin.";
RL   J. Biochem. 83:35-48(1978).
RN   [4]
RP   SEQUENCE OF 223-288 AND 572-608.
RX   MEDLINE=76260153; PubMed=956149;
RA   Isemura S., Ikenaka T.;
RT   "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
RT   the amino acid sequences of four fragments.";
RL   J. Biochem. 79:1183-1196(1976).
RN   [5]
RP   SEQUENCE OF 166-174.
RC   TISSUE=Plasma;
RX   MEDLINE=87194805; PubMed=2437111;
RA   Carraway R.E., Mitra S.P., Cochrane D.E.;
RT   "Structure of a biologically active neurotensin-related peptide
RT   obtained from pepsin-treated albumin(s).";
RL   J. Biol. Chem. 262:5968-5973(1987).
RN   [6]
RP   COPPER-BINDING.
RX   MEDLINE=79001617; PubMed=80265;
RA   Aoyagi Y., Ikenaka T., Ichida F.;
RT   "Copper(II)-binding ability of human alpha-fetoprotein.";
RL   Cancer Res. 38:3483-3486(1978).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and blood
CC       flow (Potential).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01222; CAA24532.1; -.
DR   PIR; A93872; ABRTS.
DR   HSSP; P02768; 1HK1.
DR   Rat-heart-2DPAGE; P02770; -.
DR   RGD; 2085; Alb.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Direct protein sequencing; Lipid-binding; Metal-binding;
KW   Repeat; Signal.
FT   SIGNAL        1     18
FT   PROPEP       19     24
FT   CHAIN        25    608       Serum albumin.
FT   PEPTIDE     166    174       Neurotensin-related peptide.
FT   DOMAIN       25    205       Albumin 1.
FT   DOMAIN      212    397       Albumin 2.
FT   DOMAIN      404    595       Albumin 3.
FT   METAL        27     27       Copper.
FT   DISULFID     77     86       By similarity.
FT   DISULFID     99    115       By similarity.
FT   DISULFID    114    125       By similarity.
FT   DISULFID    148    193       By similarity.
FT   DISULFID    192    201       By similarity.
FT   DISULFID    224    270       By similarity.
FT   DISULFID    269    277       By similarity.
FT   DISULFID    289    303       By similarity.
FT   DISULFID    302    313       By similarity.
FT   DISULFID    340    385       By similarity.
FT   DISULFID    384    393       By similarity.
FT   DISULFID    416    462       By similarity.
FT   DISULFID    461    472       By similarity.
FT   DISULFID    485    501       By similarity.
FT   DISULFID    500    511       By similarity.
FT   DISULFID    538    583       By similarity.
FT   DISULFID    582    591       By similarity.
FT   VARIANT     262    262       V -> L.
FT   CONFLICT    174    174       Y -> L (in Ref. 5).
SQ   SEQUENCE   608 AA;  68718 MW;  5BB497A282411AB7 CRC64;

  Query Match             86.2%;  Score 106;  DB 1;  Length 608;
  Best Local Similarity   79.2%;  Pred. No. 5.5e-07;
  Matches   19;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :|||||:|||||||||::|| |||
Db         25 EAHKSEIAHRFKDLGEQHFKGLVL 48


RESULT 11
ALBU_SHEEP
ID   ALBU_SHEEP     STANDARD;      PRT;   607 AA.
AC   P14639;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor.
GN   Name=ALB;
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=90098888; PubMed=2602160;
RA   Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;
RT   "Nucleotide and deduced amino acid sequence of sheep serum albumin.";
RL   Nucleic Acids Res. 17:10495-10495(1989).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X17055; CAA34903.1; -.
DR   PIR; S06936; ABSHS.
DR   HSSP; P02768; 1HK1.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL        1     18       By similarity.
FT   PROPEP       19     24       By similarity.
FT   CHAIN        25    607       Serum albumin.
FT   DOMAIN       25    204       Albumin 1.
FT   DOMAIN      211    396       Albumin 2.
FT   DOMAIN      403    594       Albumin 3.
FT   METAL        27     27       Copper (By similarity).
FT   DISULFID     77     86       By similarity.
FT   DISULFID     99    115       By similarity.
FT   DISULFID    114    125       By similarity.
FT   DISULFID    147    192       By similarity.
FT   DISULFID    191    200       By similarity.
FT   DISULFID    223    269       By similarity.
FT   DISULFID    268    276       By similarity.
FT   DISULFID    288    302       By similarity.
FT   DISULFID    301    312       By similarity.
FT   DISULFID    339    384       By similarity.
FT   DISULFID    383    392       By similarity.
FT   DISULFID    415    461       By similarity.
FT   DISULFID    460    471       By similarity.
FT   DISULFID    484    500       By similarity.
FT   DISULFID    499    510       By similarity.
FT   DISULFID    537    582       By similarity.
FT   DISULFID    581    590       By similarity.
SQ   SEQUENCE   607 AA;  69188 MW;  84979A87F8B86596 CRC64;

  Query Match             85.4%;  Score 105;  DB 1;  Length 607;
  Best Local Similarity   79.2%;  Pred. No. 7.7e-07;
  Matches   19;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||:|||| |||||||: |||
Db         25 DTHKSEIAHRFNDLGEENFQGLVL 48


RESULT 12
ALBU_HORSE
ID   ALBU_HORSE     STANDARD;      PRT;   607 AA.
AC   P35747;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor (Allergen Equ c 3).
GN   Name=ALB;
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC   TISSUE=Liver;
RX   MEDLINE=93345495; PubMed=8344282;
RA   Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT   "X-ray and primary structure of horse serum albumin (Equus caballus)
RT   at 0.27-nm resolution.";
RL   Eur. J. Biochem. 215:205-212(1993).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X74045; CAA52194.1; -.
DR   PIR; S34053; ABHOS.
DR   HSSP; P02768; 1HK1.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Allergen; Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL        1     18       By similarity.
FT   PROPEP       19     24       By similarity.
FT   CHAIN        25    607       Serum albumin.
FT   DOMAIN       25    204       Albumin 1.
FT   DOMAIN      211    396       Albumin 2.
FT   DOMAIN      403    594       Albumin 3.
FT   METAL        27     27       Copper (By similarity).
FT   DISULFID     77     86
FT   DISULFID     99    115
FT   DISULFID    114    125
FT   DISULFID    147    192
FT   DISULFID    191    200
FT   DISULFID    223    269
FT   DISULFID    268    276
FT   DISULFID    288    302
FT   DISULFID    301    312
FT   DISULFID    339    384
FT   DISULFID    383    392
FT   DISULFID    415    461
FT   DISULFID    460    471
FT   DISULFID    484    500
FT   DISULFID    499    510
FT   DISULFID    537    582
FT   DISULFID    581    590
SQ   SEQUENCE   607 AA;  68598 MW;  256F6E830A1B90C5 CRC64;

  Query Match             81.3%;  Score 100;  DB 1;  Length 607;
  Best Local Similarity   75.0%;  Pred. No. 4.2e-06;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||:|||| ||||::|| |||
Db         25 DTHKSEIAHRFNDLGEKHFKGLVL 48


RESULT 13
Q8CG74
ID   Q8CG74      PRELIMINARY;      PRT;   205 AA.
AC   Q8CG74;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Albumin (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvevTACfBr;
RA   Van Reeth T., Dreze P.L., Gabant P., Szpirer C., Szpirer J.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ277794; CAC81903.1; -.
DR   HSSP; P02768; 1HK1.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 1.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   SMART; SM00103; ALBUMIN; 1.
FT   NON_TER     205    205
SQ   SEQUENCE   205 AA;  23624 MW;  0A982E85C7474BF9 CRC64;

  Query Match             79.7%;  Score 98;  DB 2;  Length 205;
  Best Local Similarity   70.8%;  Pred. No. 2.7e-06;
  Matches   17;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :|||||:|||: ||||::|| |||
Db         25 EAHKSEIAHRYNDLGEQHFKGLVL 48


RESULT 14
ALBU_MOUSE
ID   ALBU_MOUSE     STANDARD;      PRT;   608 AA.
AC   P07724; Q61802;
DT   01-APR-1988 (Rel. 07, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor.
GN   Name=Alb; Synonyms=Alb-1, Alb1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Tongue;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Kidney, and Liver;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 99-516 FROM N.A.
RX   MEDLINE=88216123; PubMed=2452956;
RA   Minghetti P.P., Law S.W., Dugaiczyk A.;
RT   "The rate of molecular evolution of alpha-fetoprotein approaches that
RT   of pseudogenes.";
RL   Mol. Biol. Evol. 2:347-358(1985).
RN   [5]
RP   SEQUENCE OF 477-551 FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=90269606; PubMed=1971802; DOI=10.1016/0378-1119(90)90030-U;
RA   Boccaccio C., Deschatrette J., Meunier-Rotival M.;
RT   "Empty and occupied insertion site of the truncated LINE-1 repeat
RT   located in the mouse serum albumin-encoding gene.";
RL   Gene 88:181-186(1990).
RN   [6]
RP   SEQUENCE OF 25-44.
RC   TISSUE=Liver;
RX   MEDLINE=93162044; PubMed=1286668;
RA   Giometti C.S., Taylor J., Tollaksen S.L.;
RT   "Mouse liver protein database: a catalog of proteins detected by two-
RT   dimensional gel electrophoresis.";
RL   Electrophoresis 13:970-991(1992).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ011413; CAA09617.1; -.
DR   EMBL; AK010025; BAB26650.1; -.
DR   EMBL; BC024643; AAH24643.1; -.
DR   EMBL; BC049971; AAH49971.1; -.
```

```
DR   EMBL; M16111; AAA37190.1; -.
DR   EMBL; X13060; CAA31458.1; -.
DR   PIR; A05139; A05139.
DR   HSSP; P02768; 1HK1.
DR   SWISS-2DPAGE; P07724; MOUSE.
DR   MGD; MGI:87991; Alb1.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Direct protein sequencing; Lipid-binding; Metal-binding;
KW   Repeat; Signal.
FT   SIGNAL        1     18       By similarity.
FT   PROPEP       19     24
FT   CHAIN        25    608       Serum albumin.
FT   DOMAIN       25    205       Albumin 1.
FT   DOMAIN      212    397       Albumin 2.
FT   DOMAIN      404    595       Albumin 3.
FT   METAL        27     27       Copper.
FT   DISULFID     77     86       By similarity.
FT   DISULFID     99    115       By similarity.
FT   DISULFID    114    125       By similarity.
FT   DISULFID    148    193       By similarity.
FT   DISULFID    192    201       By similarity.
FT   DISULFID    224    270       By similarity.
FT   DISULFID    269    277       By similarity.
FT   DISULFID    289    303       By similarity.
FT   DISULFID    302    313       By similarity.
FT   DISULFID    340    385       By similarity.
FT   DISULFID    384    393       By similarity.
FT   DISULFID    416    462       By similarity.
FT   DISULFID    461    472       By similarity.
FT   DISULFID    485    501       By similarity.
FT   DISULFID    500    511       By similarity.
FT   DISULFID    538    583       By similarity.
FT   DISULFID    582    591       By similarity.
FT   CONFLICT     27     27       H -> D (in Ref. 6).
FT   CONFLICT     33     33       H -> D (in Ref. 6).
FT   CONFLICT     41     41       Q -> I (in Ref. 6).
SQ   SEQUENCE   608 AA;  68692 MW;  292F7C7EED3A61B4 CRC64;

  Query Match             79.7%;  Score 98;  DB 1;  Length 608;
  Best Local Similarity   70.8%;  Pred. No. 8.4e-06;
  Matches   17;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :|||||:|||: ||||::|| |||
Db         25 EAHKSEIAHRYNDLGEQHFKGLVL 48


RESULT 15
Q8C7H3
ID   Q8C7H3      PRELIMINARY;      PRT;   608 AA.
AC   Q8C7H3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus adult male liver tumor cDNA, RIKEN full-length enriched
DE   library, clone:C730030P03 product:albumin 1, full insert
DE   sequence.
GN   Name=Alb1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK050248; BAC34145.1; -.
DR   HSSP; P02768; 1HK1.
DR   MGD; MGI:87991; Alb1.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
SQ   SEQUENCE   608 AA;  68722 MW;  292F600EED3A61B4 CRC64;

  Query Match             79.7%;  Score 98;  DB 2;  Length 608;
  Best Local Similarity   70.8%;  Pred. No. 8.4e-06;
  Matches   17;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
```

```
              :|||||:|||: ||||::|| |||
Db         25 EAHKSEIAHRYNDLGEQHFKGLVL 48


RESULT 16
Q7YSG3
ID   Q7YSG3      PRELIMINARY;      PRT;   584 AA.
AC   Q7YSG3;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Albumin (Fragment).
GN   Name=alb;
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
RA   Rumpold H., Valenta R., Spitzauer S.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ487677; CAD32275.1; -.
DR   HSSP; P02768; 1E7B.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
FT   NON_TER       1      1
SQ   SEQUENCE   584 AA;  65908 MW;  B51002F12902C9CE CRC64;

  Query Match             78.9%;  Score 97;  DB 2;  Length 584;
  Best Local Similarity   70.8%;  Pred. No. 1.1e-05;
  Matches   17;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :||:|||:|||| |||||:|: |||
Db          1 EAHQSEIAHRFNDLGEEHFRGLVL 24


RESULT 17
ALBU_PIG
ID   ALBU_PIG       STANDARD;      PRT;   605 AA.
AC   P08835; Q29018;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor (Fragment).
GN   Name=ALB;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=89016582; PubMed=3174440;
RA   Baldwin G.S., Weinstock J.;
RT   "Nucleotide sequence of porcine liver albumin.";
RL   Nucleic Acids Res. 16:9045-9045(1988).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X12422; CAA30970.1; -.
DR   EMBL; M36787; AAA30988.1; -.
DR   PIR; S01382; ABPGS.
DR   HSSP; P02768; 1E7H.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     16       By similarity.
FT   PROPEP       17     22       By similarity.
FT   CHAIN        23    605       Serum albumin.
FT   DOMAIN       23    202       Albumin 1.
FT   DOMAIN      209    394       Albumin 2.
FT   DOMAIN      401    592       Albumin 3.
FT   METAL        31     31       Copper (By similarity).
FT   DISULFID     75     84       By similarity.
FT   DISULFID     97    113       By similarity.
FT   DISULFID    112    123       By similarity.
FT   DISULFID    145    190       By similarity.
FT   DISULFID    189    198       By similarity.
FT   DISULFID    221    267       By similarity.
FT   DISULFID    266    274       By similarity.
FT   DISULFID    286    300       By similarity.
FT   DISULFID    299    310       By similarity.
FT   DISULFID    337    382       By similarity.
FT   DISULFID    381    390       By similarity.
FT   DISULFID    413    459       By similarity.
FT   DISULFID    458    469       By similarity.
FT   DISULFID    482    498       By similarity.
FT   DISULFID    497    508       By similarity.
FT   DISULFID    535    580       By similarity.
FT   DISULFID    579    588       By similarity.
FT   CONFLICT    562    562       E -> D (in Ref. 1; AAA30988).
SQ   SEQUENCE   605 AA;  69410 MW;  3E556B0DD1A1F4FF CRC64;

  Query Match             78.9%;  Score 97;  DB 1;  Length 605;
  Best Local Similarity   75.0%;  Pred. No. 1.2e-05;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | :|||:||||||||: || |||
Db         23 DTYKSEIAHRFKDLGEQYFKGLVL 46


RESULT 18
Q68NH7
ID   Q68NH7      PRELIMINARY;      PRT;   607 AA.
AC   Q68NH7;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Albumin.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sun S., Deng J., Zhou Y., Lu J., Wu X.;
RT   "Porcine serum albumin gene.";
```

```
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY663543; AAT98610.1; -.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
SQ   SEQUENCE   607 AA;  69691 MW;  7B8DA13543CA99D8 CRC64;

  Query Match             78.9%;  Score 97;  DB 2;  Length 607;
  Best Local Similarity   75.0%;  Pred. No. 1.2e-05;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | :|||:|||||||||: || |||
Db         25 DTYKSEIAHRFKDLGEQYFKGLVL 48


RESULT 19
ALBU_FELCA
ID   ALBU_FELCA     STANDARD;      PRT;   608 AA.
AC   P49064;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor (Allergen Fel d 2).
GN   Name=ALB;
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96194824; PubMed=8647469; DOI=10.1016/0378-1119(95)00851-9;
RA   Hilger C., Grigioni F., Kohnen M., Hentges F.;
RT   "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RL   Gene 169:295-296(1996).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- ALLERGEN: Causes an allergic reaction in human.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X84842; CAA59279.1; -.
DR   PIR; JC4660; S57632.
DR   HSSP; P02768; 1E7B.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Allergen; Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL       1     18       By similarity.
FT   PROPEP      19     24       By similarity.
FT   CHAIN       25    608       Serum albumin.
FT   DOMAIN      25    205       Albumin 1.
FT   DOMAIN     212    397       Albumin 2.
FT   DOMAIN     404    595       Albumin 3.
FT   METAL       27     27       Copper.
FT   DISULFID    77     86       By similarity.
FT   DISULFID    99    115       By similarity.
FT   DISULFID   114    125       By similarity.
FT   DISULFID   148    193       By similarity.
FT   DISULFID   192    201       By similarity.
FT   DISULFID   224    270       By similarity.
FT   DISULFID   269    277       By similarity.
FT   DISULFID   289    303       By similarity.
FT   DISULFID   302    313       By similarity.
FT   DISULFID   340    385       By similarity.
FT   DISULFID   384    393       By similarity.
FT   DISULFID   416    462       By similarity.
FT   DISULFID   461    472       By similarity.
FT   DISULFID   485    501       By similarity.
FT   DISULFID   500    511       By similarity.
FT   DISULFID   538    583       By similarity.
FT   DISULFID   582    591       By similarity.
SQ   SEQUENCE   608 AA;  68659 MW;  07E629CAC5F60E5F CRC64;

  Query Match             78.9%;  Score 97;  DB 1;  Length 608;
  Best Local Similarity   70.8%;  Pred. No. 1.2e-05;
  Matches   17;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :||:||:|||| |||||:|: |||
Db         25 EAHQSEIAHRFNDLGEEHFRGLVL 48


RESULT 20
Q6WDN9
ID   Q6WDN9      PRELIMINARY;      PRT;   608 AA.
AC   Q6WDN9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Preproalbumin precursor.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Mihaylova-Todorova S., Choe S.M., Westfall D.P.;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY294645; AAQ20088.1; -.
DR   HSSP; P02768; 1AO6.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Signal.
FT   SIGNAL       1     21       Potential.
FT   CHAIN       25    608       albumin.
SQ   SEQUENCE   608 AA;  68889 MW;  BBD510A78D0261BA CRC64;

  Query Match             78.9%;  Score 97;  DB 2;  Length 608;
  Best Local Similarity   75.0%;  Pred. No. 1.2e-05;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :|||||:|||| |||| :|| |||
Db         25 EAHKSEIAHRFNDLGEGHFKGLVL 48


RESULT 21
```

```
ALBU_MERUN
ID   ALBU_MERUN     STANDARD;      PRT;   609 AA.
AC   O35090;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor.
GN   Name=ALB;
OS   Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC   Meriones.
OX   NCBI_TaxID=10047;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGS IDR; TISSUE=Liver;
RX   MEDLINE=98116663; PubMed=9455485;
RA   Yoshida K., Seto-Ohshima A., Sinohara H.;
RT   "Sequencing of cDNA encoding serum albumin and its extrahepatic
RT   synthesis in the Mongolian gerbil, Meriones unguiculatus.";
RL   DNA Res. 4:351-354(1997).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB006197; BAA21765.1; -.
DR   PIR; JC5838; JC5838.
DR   HSSP; P02768; 1E7B.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL        1     18       By similarity.
FT   PROPEP       19     24       By similarity.
FT   CHAIN        25    609       Serum albumin.
FT   DOMAIN       25    206       Albumin 1.
FT   DOMAIN      213    398       Albumin 2.
FT   DOMAIN      405    596       Albumin 3.
FT   METAL        28     28       Copper.
FT   DISULFID     78     87       By similarity.
FT   DISULFID    100    116       By similarity.
FT   DISULFID    115    126       By similarity.
FT   DISULFID    149    194       By similarity.
FT   DISULFID    193    202       By similarity.
FT   DISULFID    225    271       By similarity.
FT   DISULFID    270    278       By similarity.
FT   DISULFID    290    304       By similarity.
FT   DISULFID    303    314       By similarity.
FT   DISULFID    341    386       By similarity.
FT   DISULFID    385    394       By similarity.
FT   DISULFID    417    463       By similarity.
FT   DISULFID    462    473       By similarity.
FT   DISULFID    486    502       By similarity.
FT   DISULFID    501    512       By similarity.
FT   DISULFID    539    584       By similarity.
FT   DISULFID    583    592       By similarity.
SQ   SEQUENCE   609 AA;  68940 MW;  9CA5F97F67EF1A48 CRC64;

  Query Match             78.9%;  Score 97;  DB 1;  Length 609;
  Best Local Similarity   78.3%;  Pred. No. 1.2e-05;
  Matches   18;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AHKSEVAHRFKDLGEENFKALVL 24
              |||||:|||:|||||: || |||
Db         27 AHKSEIAHRYKDLGEKYFKGLVL 49


RESULT 22
Q9TRA5
ID   Q9TRA5      PRELIMINARY;      PRT;    40 AA.
AC   Q9TRA5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Antagonist protein (Fragment).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94359514; PubMed=8078511;
RA   Xu A., Narayanan N.;
RT   "Purification, amino-terminal sequence and functional properties of a
RT   64 kDa cytosolic protein from heart muscle capable of modulating
RT   calcium transport across the sarcoplasmic reticulum in vitro.";
RL   Mol. Cell. Biochem. 132:7-14(1994).
DR   HSSP; P02768; 1N5U.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 1.
SQ   SEQUENCE   40 AA;  4694 MW;  BA7E0B69C6CE858C CRC64;

  Query Match             75.6%;  Score 93;  DB 2;  Length 40;
  Best Local Similarity   70.8%;  Pred. No. 2.6e-06;
  Matches   17;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :|||||:|||| |:|||:|  |||
Db          1 EAHKSEIAHRFNDVGEEHFIGLVL 24


RESULT 23
ALBU_RABIT
ID   ALBU_RABIT     STANDARD;      PRT;   608 AA.
AC   P49065;
DT   01-FEB-1996 (Rel. 33, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor.
GN   Name=ALB;
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white; TISSUE=Liver;
RX   MEDLINE=97275135; PubMed=9129029;
RA   Syed S., Schuyler P.D., Kulczycky M., Sheffield W.P.;
RT   "Potent antithrombin activity and delayed clearance from the
RT   circulation characterize recombinant hirudin genetically fused to
RT   albumin.";
RL   Blood 89:3243-3252(1997).
RN   [2]
RP   REVISIONS TO 322-323 AND 506-507.
RA   Sheffield W.P.;
```

```
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U18344; AAB58347.2; -.
DR   HSSP; P02768; 1E7B.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
KW   Copper; Lipid-binding; Metal-binding; Repeat; Signal.
FT   SIGNAL        1     18       By similarity.
FT   PROPEP       19     24       By similarity.
FT   CHAIN        25    608       Serum albumin.
FT   DOMAIN       25    205       Albumin 1.
FT   DOMAIN      212    397       Albumin 2.
FT   DOMAIN      404    595       Albumin 3.
FT   METAL        27     27       Copper.
FT   DISULFID     77     86       By similarity.
FT   DISULFID     99    115       By similarity.
FT   DISULFID    114    125       By similarity.
FT   DISULFID    148    193       By similarity.
FT   DISULFID    192    201       By similarity.
FT   DISULFID    224    270       By similarity.
FT   DISULFID    269    277       By similarity.
FT   DISULFID    289    303       By similarity.
FT   DISULFID    302    313       By similarity.
FT   DISULFID    340    385       By similarity.
FT   DISULFID    384    393       By similarity.
FT   DISULFID    416    462       By similarity.
FT   DISULFID    461    472       By similarity.
FT   DISULFID    485    501       By similarity.
FT   DISULFID    500    511       By similarity.
FT   DISULFID    538    583       By similarity.
FT   DISULFID    582    591       By similarity.
SQ   SEQUENCE   608 AA;  68909 MW;  9EECAFDA86B1EF09 CRC64;

  Query Match             75.6%;  Score 93;  DB 1;  Length 608;
  Best Local Similarity   70.8%;  Pred. No. 4.6e-05;
  Matches   17;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :|||||:|||| |:|||:|  |||
Db         25 EAHKSEIAHRFNDVGEEHFIGLVL 48


RESULT 24
Q9QUX8
ID   Q9QUX8      PRELIMINARY;      PRT;    20 AA.
AC   Q9QUX8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   70 kDa seizure activity-linked albumin-like glycoprotein (Fragment).
OS   Rattus sp. .
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96087830; PubMed=8587647;
RA   Onozuka M., Imai S., Isobe T., Yen C.T., Watanabe K.;
RT   "Purification and characterization of a novel 70-kDa brain protein
RT   associated with seizure activities.";
RL   Neurochem. Res. 20:901-905(1995).
DR   HSSP; P02768; 1N5U.
SQ   SEQUENCE   20 AA;  2381 MW;  534A23C0F9F70F7D CRC64;

  Query Match             74.8%;  Score 92;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 1.8e-06;
  Matches   16;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFK 20
              :|||||:|||||||||: ||
Db          1 EAHKSEIAHRFKDLGEQXFK 20


RESULT 25
ALBU_CANFA
ID   ALBU_CANFA     STANDARD;      PRT;   608 AA.
AC   P49822; O77705; Q9TSZ4;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor (Allergen Can f 3).
GN   Name=ALB;
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Beagle; TISSUE=Liver;
RA   Hilger C.;
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=20148667; PubMed=10669848; DOI=10.1016/S0091-6749(00)90077-0;
RA   Pandjaitan B., Swoboda I., Brandejsky-Pichler F., Rumpold H.,
RA   Valenta R., Spitzauer S.;
RT   "Escherichia coli expression and purification of recombinant dog
RT   albumin, a cross-reactive animal allergen.";
RL   J. Allergy Clin. Immunol. 105:279-285(2000).
RN   [3]
RP   SEQUENCE OF 25-48.
RX   MEDLINE=75011422; PubMed=4414013;
RA   Dixon J.W., Sarkar B.;
RT   "Isolation, amino acid sequence and copper(II)-binding properties of
RT   peptide (1-24) of dog serum albumin.";
RL   J. Biol. Chem. 249:5872-5877(1974).
RN   [4]
RP   SEQUENCE OF 25-38.
RC   TISSUE=Heart;
RX   MEDLINE=98163340; PubMed=9504812;
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
RN   [5]
RP   SEQUENCE OF 215-478 FROM N.A.
RC   TISSUE=Salivary gland;
RX   MEDLINE=94201492; PubMed=7512102;
RA   Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,
RA   Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
RT   "Molecular characterization of dog albumin as a cross-reactive
RT   allergen.";
RL   J. Allergy Clin. Immunol. 93:614-627(1994).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
```

```
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC    -!- SUBCELLULAR LOCATION: Secreted.
CC    -!- TISSUE SPECIFICITY: Plasma.
CC    -!- ALLERGEN: Causes an allergic reaction in human.
CC    -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC    -!- SIMILARITY: Contains 3 albumin domains.
CC    ---------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------
DR    EMBL; AJ133489; CAB64867.1; -.
DR    EMBL; Y17737; CAA76841.1; -.
DR    EMBL; S72946; AAB30434.1; -.
DR    HSSP; P02768; 1E7B.
DR    HSC-2DPAGE; P49822; DOG.
DR    InterPro; IPR000264; Serum_albumin.
DR    Pfam; PF00273; Serum_albumin; 3.
DR    PRINTS; PR00802; SERUMALBUMIN.
DR    ProDom; PD002486; Serum_albumin; 1.
DR    SMART; SM00103; ALBUMIN; 3.
DR    PROSITE; PS00212; ALBUMIN; 3.
KW    Allergen; Copper; Direct protein sequencing; Lipid-binding;
KW    Metal-binding; Repeat; Signal.
FT    SIGNAL        1     18       Potential.
FT    PROPEP       19     24
FT    CHAIN        25    608       Serum albumin.
FT    DOMAIN       25    205       Albumin 1.
FT    DOMAIN      212    397       Albumin 2.
FT    DOMAIN      404    595       Albumin 3.
FT    METAL        27     27       Copper (By similarity).
FT    DISULFID     77     86       By similarity.
FT    DISULFID     99    115       By similarity.
FT    DISULFID    114    125       By similarity.
FT    DISULFID    148    193       By similarity.
FT    DISULFID    192    201       By similarity.
FT    DISULFID    224    270       By similarity.
FT    DISULFID    269    277       By similarity.
FT    DISULFID    289    303       By similarity.
FT    DISULFID    302    313       By similarity.
FT    DISULFID    340    385       By similarity.
FT    DISULFID    384    393       By similarity.
FT    DISULFID    416    462       By similarity.
FT    DISULFID    461    472       By similarity.
FT    DISULFID    485    501       By similarity.
FT    DISULFID    500    511       By similarity.
FT    DISULFID    538    583       By similarity.
FT    DISULFID    582    591       By similarity.
FT    CONFLICT      1     26       MKWVTFISLFFLFSSAYSRGLVRREA -> MDT (in
FT                                  Ref. 2).
FT    CONFLICT    146    146       A -> R (in Ref. 2).
FT    CONFLICT    206    206       I -> T (in Ref. 2).
FT    CONFLICT    349    349       V -> A (in Ref. 2).
FT    CONFLICT    359    359       S -> A (in Ref. 1).
FT    CONFLICT    448    448       V -> VV (in Ref. 5).
FT    CONFLICT    474    474       D -> E (in Ref. 1).
SQ    SEQUENCE   608 AA;  68606 MW;  3CF1C8FF7DD8FC06 CRC64;

  Query Match             74.8%;  Score 92;  DB 1;  Length 608;
  Best Local Similarity   66.7%;  Pred. No. 6.5e-05;
  Matches   16;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :|:|||:|||: |||||:|: |||
Db         25 EAYKSEIAHRYNDLGEEHFRGLVL 48


RESULT 26
Q7M3A0
ID   Q7M3A0      PRELIMINARY;      PRT;    30 AA.
AC   Q7M3A0;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Serum albumin, milk-derived (Fragment).
OS   Tachyglossus aculeatus (Australian echidna).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Monotremata; Tachyglossidae; Tachyglossus.
OX   NCBI_TaxID=9261;
RN   [1]
RP   SEQUENCE.
RA   Teahan C.G., McKenzie H.A., Griffiths M.;
RT   "Some monotreme milk "whey" and blood proteins.";
RL   Comp. Biochem. Physiol. 99:99-118(1991).
DR   PIR; B61511; B61511.
DR   HSSP; P02768; 1BKE.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 1.
FT   NON_TER       1      1
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3482 MW;  01AEA775ECCE474C CRC64;

  Query Match             72.4%;  Score 89;  DB 2;  Length 30;
  Best Local Similarity   66.7%;  Pred. No. 7.5e-06;
  Matches   16;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              || |||: ||:|:|||::|||| |
Db          1 DAQKSELGHRYKELGEDHFKALAL 24


RESULT 27
Q9R4X7
ID   Q9R4X7      PRELIMINARY;      PRT;    20 AA.
AC   Q9R4X7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   66 kDa SEROREACTIVE protein/serum albumin homolog (Fragment).
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94343500; PubMed=8064836;
RA   Deshpande R.G., Khan M.B., Bhat D.A., Navalkar R.G.;
RT   "Purification and partial characterisation of a novel 66-kDa
RT   seroreactive protein of Mycobacterium tuberculosis H37Rv.";
RL   J. Med. Microbiol. 41:173-178(1994).
SQ   SEQUENCE   20 AA;  2393 MW;  534A232072870F7D CRC64;

  Query Match             70.7%;  Score 87;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 9.7e-06;
  Matches   16;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFK 20
              | ||||:|||||||||  ||
Db          1 DKHKSEIAHRFKDLGEXPFK 20


RESULT 28
Q9TRW8
ID   Q9TRW8      PRELIMINARY;      PRT;    25 AA.
AC   Q9TRW8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

```
DE   Serum albumin (Fragment).
OS   Trichosurus vulpecula (Brush-tailed possum).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
OX   NCBI_TaxID=9337;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=91330574; PubMed=1868684;
RA   Grigor M.R., Bennett B.L., Carne A., Cowan P.E.;
RT   "Whey proteins of the common brushtail possum (Trichosurus vulpecula):
RT   isolation, characterization and changes in concentration in milk
RT   during lactation of transferrin, alpha-lactalbumin and serum
RT   albumin.";
RL   Comp. Biochem. Physiol. 98B:451-459(1991).
DR   HSSP; P02768; 1O9X.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 1.
FT   NON_TER       1      1
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2813 MW;  FD4D08B2260C67F2 CRC64;

  Query Match             70.7%;  Score 87;  DB 2;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 1.2e-05;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              || ||||| |::|||:| |||||
Db          1 DAPKSEVAKRYRDLGEKNVKALVL 24


RESULT 29
Q6B3Z0
ID   Q6B3Z0      PRELIMINARY;      PRT;   583 AA.
AC   Q6B3Z0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Serum albumin (Fragment).
GN   Name=ESA;
OS   Elephas maximus (Indian elephant).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Proboscidea; Elephantidae; Elephas.
OX   NCBI_TaxID=9783;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lazar J., Rasmussen B., Greenwood D.R., Bang I.-S., Prestwich G.D.;
RT   "Elephant Albumin: A Multi-purpose Pheromone Shuttle.";
RL   Chem. Biol. 0:0-0(2004).
DR   EMBL; AY684122; AAT90502.1; -.
DR   GO; GO:0005615; C:extracellular space; IEA.
DR   GO; GO:0005386; F:carrier activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
FT   NON_TER       1      1
SQ   SEQUENCE   583 AA;  65767 MW;  4EC031C2EC8B9141 CRC64;

  Query Match             68.3%;  Score 84;  DB 2;  Length 583;
  Best Local Similarity   68.2%;  Pred. No. 0.00094;
  Matches   15;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKALVL 24
              :|||:|||:|||||: || |:|
Db          1 YKSEIAHRYKDLGEDLFKGLLL 22


RESULT 30
ALBU_CHICK
ID   ALBU_CHICK     STANDARD;      PRT;   615 AA.
AC   P19121;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).
GN   Name=ALB;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Cassady A.I., Salklld C.K., Baverstock P., Wallace J.C.;
RL   Submitted (JUL-1991) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-28 FROM N.A.
RX   MEDLINE=83161037; PubMed=6187737;
RA   Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.;
RT   "The 5' noncoding and flanking regions of the avian very low density
RT   apolipoprotein II and serum albumin genes. Homologies with the egg
RT   white protein genes.";
RL   J. Biol. Chem. 258:4556-4564(1983).
RN   [3]
RP   SEQUENCE OF 19-30.
RX   MEDLINE=78019943; PubMed=911327;
RA   Rosen A.M., Geller D.M.;
RT   "Chicken microsomal albumin: amino terminal sequence of chicken
RT   proalbumin.";
RL   Biochem. Biophys. Res. Commun. 78:1060-1066(1977).
RN   [4]
RP   ALLERGENIC PROPERTIES.
RX   MEDLINE=21381307; PubMed=11488669;
RA   Quirce S., Maranon F., Umpierrez A., de las Heras M.,
RA   Fernandez-Caldas E., Sastre J.;
RT   "Chicken serum albumin (Gal d 5*) is a partially heat-labile inhalant
RT   and food allergen implicated in the bird-egg syndrome.";
RL   Allergy 56:754-762(2001).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.
CC       Partially heat-labile allergen that may cause both respiratory and
CC       food-allergy symptoms in patients with the bird-egg syndrome.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains 3 albumin domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X60688; CAA43098.1; -.
DR   EMBL; V00381; CAA23680.1; -.
DR   PIR; S15571; ABCHS.
DR   HSSP; P02768; 1HK1.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
```

```
KW   Allergen; Copper; Direct protein sequencing; Lipid-binding;
KW   Metal-binding; Repeat; Signal.
FT   SIGNAL       1     18
FT   PROPEP      19     23
FT   CHAIN       24    615       Serum albumin.
FT   DOMAIN      24    209       Albumin 1.
FT   DOMAIN     216    401       Albumin 2.
FT   DOMAIN     408    599       Albumin 3.
FT   METAL       30     30       Copper (By similarity).
FT   DISULFID    80     89       By similarity.
FT   DISULFID   102    118       By similarity.
FT   DISULFID   117    128       By similarity.
FT   DISULFID   152    197       By similarity.
FT   DISULFID   196    205       By similarity.
FT   DISULFID   228    274       By similarity.
FT   DISULFID   273    281       By similarity.
FT   DISULFID   293    307       By similarity.
FT   DISULFID   306    317       By similarity.
FT   DISULFID   344    389       By similarity.
FT   DISULFID   388    397       By similarity.
FT   DISULFID   420    466       By similarity.
FT   DISULFID   465    476       By similarity.
FT   DISULFID   489    505       By similarity.
FT   DISULFID   504    515       By similarity.
FT   DISULFID   542    587       By similarity.
FT   DISULFID   586    595       By similarity.
FT   CARBOHYD   500    500       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT    24     24       F -> M (in Ref. 3).
SQ   SEQUENCE   615 AA;  69918 MW;  E59E4BBCAEC066C6 CRC64;

  Query Match             65.9%;  Score 81;  DB 1;  Length 615;
  Best Local Similarity   63.6%;  Pred. No. 0.0028;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKALVL 24
              ||||:|||: || || |||: :
Db         30 HKSEIAHRYNDLKEETFKAVAM 51


RESULT 31
ALB1_TRASC
ID   ALB1_TRASC     STANDARD;      PRT;    40 AA.
AC   P81188;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   67 kDa serum albumin (Alb-1) (Fragment).
OS   Trachemys scripta (Red-eared slider turtle) (Pseudemys scripta).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Testudines; Cryptodira; Testudinoidea; Emydidae; Trachemys.
OX   NCBI_TaxID=34903;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=98103404; PubMed=9440230; DOI=10.1016/S0305-0491(97)00060-6;
RA   Brown M.A., Chambers G.K., Licht P.;
RT   "Purification and partial amino acid sequences of two distinct
RT   albumins from turtle plasma.";
RL   Comp. Biochem. Physiol. 118B:367-374(1997).
RN   [2]
RP   SEQUENCE OF 1-8.
RX   MEDLINE=95309661; PubMed=7789749; DOI=10.1006/gcen.1995.1034;
RA   Selcer K.W., Palmer B.D.;
RT   "Estrogen downregulation of albumin and a 170-kDa serum protein in the
RT   turtle, Trachemys scripta.";
RL   Gen. Comp. Endocrinol. 97:340-352(1995).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- MISCELLANEOUS: In the red-eared slider turtle, there are two forms
CC       of albumin, ALB-1 and ALB-2.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
DR   HSSP; P02768; 1E7H.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 1.
DR   PROSITE; PS00212; ALBUMIN; PARTIAL.
KW   Copper; Direct protein sequencing; Lipid-binding; Metal-binding.
FT   METAL        4      4       Copper (By similarity).
FT   NON_TER     40     40
SQ   SEQUENCE   40 AA;  4682 MW;  5FAC9E49E2789BB0 CRC64;

  Query Match             61.8%;  Score 76;  DB 1;  Length 40;
  Best Local Similarity   63.6%;  Pred. No. 0.00085;
  Matches   14;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKALVL 24
              ||||: ||| || || ||    |
Db          4 HKSEIVHRFNDLKEEKFKGAAL 25


RESULT 32
Q9TR98
ID   Q9TR98      PRELIMINARY;      PRT;    17 AA.
AC   Q9TR98;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Albumin (Fragment).
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE.
RA   Miller M.J., Parmelee D.C., Benjamin T., Sechi S., Dooley K.L.,
RA   Kadlubar F.F.;
RT   "Plasma proteins as early biomarkers of exposure to carcinogenic
RT   aromatic amines.";
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
SQ   SEQUENCE   17 AA;  2024 MW;  1D39F70F7D23B269 CRC64;

  Query Match             57.7%;  Score 71;  DB 2;  Length 17;
  Best Local Similarity   70.6%;  Pred. No. 0.0019;
  Matches   12;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEE 17
              :|:|||:|||: |||||
Db          1 EAYKSEIAHRYNDLGEE 17


RESULT 33
Q9PRW0
ID   Q9PRW0      PRELIMINARY;      PRT;    30 AA.
AC   Q9PRW0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Alpha 1-proteinase inhibitor (Fragment).
OS   Struthio camelus (Ostrich).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;
OC   Struthio.
OX   NCBI_TaxID=8801;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94341467; PubMed=8063009; DOI=10.1016/0020-711X(94)90114-7;
RA   Kuhn C.R., Naude R.J., Travis J., Oelofsen W.;
RT   "The isolation and partial characterization of alpha 1-proteinase
RT   inhibitor from the serum of the ostrich (Struthio camelus).";
RL   Int. J. Biochem. 26:843-853(1994).
SQ   SEQUENCE   30 AA;  3557 MW;  7775AA786EE30AC2 CRC64;

  Query Match             57.7%;  Score 71;  DB 2;  Length 30;
  Best Local Similarity   59.1%;  Pred. No. 0.0035;
```

```
  Matches   13;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKALVL 24
              ||||:|||: || ||  ||: :
Db          4 HKSEIAHRYNDLKEETDKAVXM 25


RESULT 34
Q8E1L1
ID   Q8E1L1      PRELIMINARY;      PRT;   144 AA.
AC   Q8E1L1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Transcriptional regulator, MarR family.
GN   OrderedLocusNames=SAG0343;
OS   Streptococcus agalactiae (serotype V).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=216466;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2603 V/R / Serotype V;
RX   MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
RA   Tettelin H., Masignani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,
RA   Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA   Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA   DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA   Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
RA   Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA   Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,
RA   Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA   Fraser C.M.;
RT   "Complete genome sequence and comparative genomic analysis of an
RT   emerging human pathogen, serotype V Streptococcus agalactiae.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
CC   -!- SIMILARITY: Contains 1 HTH marR-type DNA-binding domain.
DR   EMBL; AE014207; AAM99249.1; -.
DR   TIGR; SAG0343; -.
DR   GO; GO:0005622; C:intracellular; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR000835; HTH_MarR.
DR   InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR   Pfam; PF01047; MarR; 1.
DR   PRINTS; PR00598; HTHMARR.
KW   Complete proteome; DNA-binding; Transcription;
KW   Transcription regulation.
SQ   SEQUENCE   144 AA;  16656 MW;  465EAE2AB8F8DB34 CRC64;

  Query Match             49.6%;  Score 61;  DB 2;  Length 144;
  Best Local Similarity   60.0%;  Pred. No. 0.55;
  Matches   12;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKAL 22
              ||| |:|  :|||||: | |
Db        110 HKSMVSHIIEDLGEEDIKML 129


RESULT 35
Q8E728
ID   Q8E728      PRELIMINARY;      PRT;   144 AA.
AC   Q8E728;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein gbs0330.
GN   OrderedLocusNames=gbs0330;
OS   Streptococcus agalactiae (serotype III).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=216495;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NEM316 / Serotype III;
RX   MEDLINE=22242508; PubMed=12354221;
RA   Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA   Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
RA   Kunst F.;
RT   "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT   invasive neonatal disease.";
RL   Mol. Microbiol. 45:1499-1513(2002).
CC   -!- SIMILARITY: Contains 1 HTH marR-type DNA-binding domain.
DR   EMBL; AL766844; CAD45975.1; -.
DR   SagaList; gbs0330; -.
DR   GO; GO:0005622; C:intracellular; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR000835; HTH_MarR.
DR   InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR   Pfam; PF01047; MarR; 1.
DR   PRINTS; PR00598; HTHMARR.
DR   SMART; SM00347; HTH_MARR; 1.
KW   Complete proteome; DNA-binding; Transcription;
KW   Transcription regulation.
SQ   SEQUENCE   144 AA;  16656 MW;  465EAE2AB8F8DB34 CRC64;

  Query Match             49.6%;  Score 61;  DB 2;  Length 144;
  Best Local Similarity   60.0%;  Pred. No. 0.55;
  Matches   12;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKAL 22
              ||| |:|  :|||||: | |
Db        110 HKSMVSHIIEDLGEEDIKML 129


RESULT 36
NF41_NAEFO
ID   NF41_NAEFO     STANDARD;      PRT;    15 AA.
AC   P83729;
DT   29-MAR-2004 (Rel. 43, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Unknown protein NF041 from 2D-PAGE (Fragment).
OS   Naegleria fowleri.
OC   Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
OX   NCBI_TaxID=5763;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=Nf 66;
RA   Omura M., Furushima-Shimogawara R., Izumiyama S., Endo T.;
RT   "Comparative study of protein profiles on pathogenic and nonpathogenic
RT   Naegleria species by 2D-PAGE.";
RL   J. Eukaryot. Microbiol. 0:0-0(2004).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 5.9, its MW is: 47.0 kDa.
KW   Direct protein sequencing.
FT   NON_TER      15      15
SQ   SEQUENCE   15 AA;  1704 MW;  C70F7D308AEC51B9 CRC64;

  Query Match             48.8%;  Score 60;  DB 1;  Length 15;
  Best Local Similarity   73.3%;  Pred. No. 0.071;
  Matches   11;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLG 15
              | ||||:|||  |||
Db          1 DTHKSEIAHRQPDLG 15


RESULT 37
Q7M8S9
ID   Q7M8S9      PRELIMINARY;      PRT;   548 AA.
AC   Q7M8S9;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

```
DE   PUTATIVE TWO-COMPONENT SENSOR.
GN   OrderedLocusNames=WS1436;
OS   Wolinella succinogenes.
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Helicobacteraceae; Wolinella.
OX   NCBI_TaxID=844;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSMZ 1740;
RX   MEDLINE=22882897; PubMed=14500908; DOI=10.1073/pnas.1932838100;
RA   Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O.,
RA   Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,
RA   Meyer F., Lederer H., Schuster S.C.;
RT   "Complete genome sequence and analysis of Wolinella succinogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).
DR   EMBL; BX571660; CAE10497.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR004358; Bact_sens_pr_C.
DR   InterPro; IPR005467; His_kinase.
DR   InterPro; IPR003660; His_kin_HAMP.
DR   InterPro; IPR009082; His_kin_homodim.
DR   Pfam; PF00672; HAMP; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00344; BCTRLSENSOR.
DR   PROSITE; PS50109; HIS_KIN; 1.
KW   Complete proteome.
SQ   SEQUENCE   548 AA;  63212 MW;  D88DB41F238B8B8D CRC64;

  Query Match             48.0%;  Score 59;  DB 2;  Length 548;
  Best Local Similarity   66.7%;  Pred. No. 4.4;
  Matches   12;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 EVAHRFKDLGEENFKALV 23
              :| | |||| || ||| |
Db        356 QVKHHFKDLTEETFKAQV 373


RESULT 38
ALBU_NEOFS
ID   ALBU_NEOFS     STANDARD;      PRT;   101 AA.
AC   P83517;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serum albumin (Fragments).
OS   Neoceratodus forsteri (Australian lungfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Dipnoi; Ceratodontiformes; Ceratodontidae; Neoceratodus.
OX   NCBI_TaxID=7892;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RA   Metcalf V., Brennan S., George P.;
RT   "Using serum albumin to infer vertebrate phylogenies.";
RL   Appl. Bioinformatics 2:S97-S107(2003).
CC   -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC       binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC       hormones, bilirubin and drugs. Its main function is the regulation
CC       of the colloidal osmotic pressure of blood.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC   -!- SIMILARITY: Contains at least 2 albumin domains.
DR   InterPro; IPR000264; Serum_albumin.
DR   PROSITE; PS00212; ALBUMIN; PARTIAL.
KW   Copper; Direct protein sequencing; Lipid-binding; Metal-binding;
KW   Repeat.
FT   DOMAIN        1    >80       Albumin 1.
FT   METAL         4      4       Copper (By similarity).
FT   NON_CONS     28     29
FT   NON_CONS     80     81
FT   DOMAIN      <81   >101       Albumin 2.
FT   NON_TER     101    101
SQ   SEQUENCE   101 AA;  11412 MW;  A51669C76226CC43 CRC64;

  Query Match             47.2%;  Score 58;  DB 1;  Length 101;
  Best Local Similarity   45.5%;  Pred. No. 1;
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKALVL 24
              ||| :   |: :||| || ::|
Db          4 HKSNICKHFQVVGEEKFKNIIL 25


RESULT 39
Q8C7C7
ID   Q8C7C7      PRELIMINARY;      PRT;   576 AA.
AC   Q8C7C7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-
DE   length enriched library, clone:C920028B14 product:albumin 1, full
DE   insert sequence. (Fragment).
GN   Name=Alb1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
```

```
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK050644; BAC34360.1; -.
DR   HSSP; P02768; 1HK1.
DR   MGD; MGI:87991; Alb1.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001703; Alphafetoprot.
DR   InterPro; IPR000264; Serum_albumin.
DR   Pfam; PF00273; Serum_albumin; 3.
DR   PRINTS; PR00803; AFETOPROTEIN.
DR   PRINTS; PR00802; SERUMALBUMIN.
DR   ProDom; PD002486; Serum_albumin; 1.
DR   SMART; SM00103; ALBUMIN; 3.
DR   PROSITE; PS00212; ALBUMIN; 3.
FT   NON_TER       1      1
SQ   SEQUENCE   576 AA;  65002 MW;  F85733E99AE37F04 CRC64;

  Query Match             45.5%;  Score 56;  DB 2;  Length 576;
  Best Local Similarity   62.5%;  Pred. No. 13;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          9 HRFKDLGEENFKALVL 24
              :|: ||||::|| |||
Db          1 NRYNDLGEQHFKGLVL 16


RESULT 40
Q6ZQ53
ID   Q6ZQ53      PRELIMINARY;      PRT;   569 AA.
AC   Q6ZQ53;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MKIAA0765 protein (Fragment).
GN   Name=mKIAA0765;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryonic tail;
RX   PubMed=14621295;
RA   Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA   Saga Y., Nagase T., Ohara O., Koga H.;
RT   "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT   III. the complete nucleotide sequences of 500 mouse KIAA-homologous
RT   cDNAs identified by screening of terminal sequences of cDNA clones
RT   randomly sampled from size-fractionated libraries.";
RL   DNA Res. 10:167-180(2003).
DR   EMBL; AK129208; BAC98018.1; -.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00076; RRM_1; 2.
DR   SMART; SM00360; RRM; 3.
DR   PROSITE; PS50102; RRM; 2.
FT   NON_TER       1      1
SQ   SEQUENCE   569 AA;  61568 MW;  57480E18D57ED6BA CRC64;

  Query Match             43.1%;  Score 53;  DB 2;  Length 569;
  Best Local Similarity   56.5%;  Pred. No. 35;
  Matches   13;  Conservative    1;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALV 23
              |||||| || |  | | |  :|
Db        508 DAHKSEHLHRKKLNGREAFVHIV 530


Search completed: August 19, 2005, 10:55:34
Job time : 58 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 10:49:44 ; Search time 16 Seconds
(without alignments)
144.325 Million cell updates/sec

Title: US-09-846-328B-1_COPY_2_25
Perfect score: 123
Sequence: 1 DAHKSEVAHRFKDLGEENFKALVL 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	123	100.0	609	1	ABHUS	serum albumin prec
2	110	89.4	600	2	A47391	serum albumin prec
3	108	87.8	607	1	ABBOS	serum albumin prec
4	106	86.2	608	1	ABRTS	serum albumin prec
5	105	85.4	607	1	ABSHS	serum albumin prec
6	100	81.3	607	1	ABHOS	serum albumin prec
7	97	78.9	605	1	ABPGS	serum albumin prec
8	97	78.9	608	2	S57632	serum albumin prec
9	97	78.9	609	2	JC5838	albumin - Mongolia
10	92	74.8	24	2	S29749	serum albumin - do
11	89	72.4	30	2	B61511	serum albumin, mil
12	81	65.9	615	1	ABCHS	serum albumin prec
13	50	40.7	67	2	T12251	sucrose synthase (
14	50	40.7	384	2	G85895	hypothetical prote
15	50	40.7	384	2	C91051	hypothetical prote
16	50	40.7	384	2	D65028	hypothetical 43.1
17	49	39.8	153	2	A53361	thyroxine-binding
18	49	39.8	905	2	S56295	SAP155 protein - y
19	48	39.0	184	2	T20006	hypothetical prote
20	48	39.0	222	2	A97575	hypothetical prote
21	48	39.0	222	2	AI2795	conserved hypothet
22	48	39.0	474	1	VYHUD	vitamin D-binding
23	47.5	38.6	629	2	T06675	hypothetical prote
24	47	38.2	189	2	AB0230	probable exported
25	47	38.2	313	2	E83095	conserved hypothet
26	47	38.2	719	2	F96577	hypothetical prote
27	46.5	37.8	451	2	F89130	protein F52E1.10 [
28	46	37.4	180	2	T44944	hypothetical prote
29	46	37.4	220	2	B72219	ribulose-phosphate
30	45	36.6	79	2	G95324	probable TraD conj
31	45	36.6	264	2	T36873	hypothetical prote
32	45	36.6	314	2	S56055	hypothetical prote
33	45	36.6	383	2	G64667	NA+/H+ antiporter
34	45	36.6	383	2	H71848	probable na+/h+ an
35	45	36.6	388	2	AD0822	conserved hypothet
36	45	36.6	394	2	S41952	UTP-glucose glucos
37	45	36.6	441	2	A45565	cysteine proteinas
38	45	36.6	639	2	A97208	carbon monoxide de
39	45	36.6	814	2	C82860	DNA gyrase subunit
40	45	36.6	2605	2	T18552	saframycin Mx1 syn
41	44	35.8	227	2	A48412	hypothetical prote
42	44	35.8	365	2	A49235	outer membrane lip
43	44	35.8	807	2	S29242	sucrose synthase (
44	44	35.8	1161	2	T18642	hypothetical prote
45	44	35.8	1417	2	H83132	probable sensor/re

ALIGNMENTS

RESULT 1
ABHUS
serum albumin precursor [validated] - human
N;Alternate names: preproalbumin
N;Contains: kinetensin
C;Species: Homo sapiens (man)
C;Date: 29-Jul-1981 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
C;Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S36
R;Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seeburg
Nucleic Acids Res. 9, 6103-6114, 1981
A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli
A;Reference number: A93743; MUID:82081882; PMID:6171778
A;Accession: A93743
A;Molecule type: mRNA
A;Residues: 1-419,'K',421-609 <LAW>
A;Cross-references: UNIPROT:P02768; EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:g28
R;Dugaiczyk, A.; Law, S.W.; Dennison, O.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A;Reference number: A93936; MUID:82105994; PMID:6275391
A;Accession: A93936
A;Molecule type: mRNA
A;Residues: 1-120,'G',122-609 <DUG>
A;Cross-references: EMBL:V00494; NID:g28589; PIDN:CAA23753.1; PID:g28590
R;Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A;Reference number: I39427; MUID:86140099; PMID:2419329
A;Accession: I39427
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-26 <URA>
A;Cross-references: GB:M13075; NID:g178330; PIDN:AAA51688.1; PID:g553173
R;Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.
A;Reference number: I59286; MUID:94181575; PMID:8134387
A;Accession: I59286
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 282-290,'KSRFDLQ' <WAT>
A;Cross-references: GB:S69192; NID:g546032; PIDN:AAB30282.1; PID:g546033
A;Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia
R;Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-
A;Reference number: I59313; MUID:94294404; PMID:8022807
A;Accession: I59313
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 589-590,'ALPRRVKNLLLQVKLP' <MAD>
A;Cross-references: GB:S70799; NID:g547231; PIDN:AAB31177.1; PID:g547232

A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants
R;Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A;Reference number: G08292
A;Accession: G01747
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-120,'G',122-455 <MEN>
A;Cross-references: EMBL:U22961; NID:g763428; PIDN:AAA64922.1; PID:g763431
R;Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
A;Reference number: S55314; MUID:95275251; PMID:7755581
A;Accession: S55314
A;Molecule type: protein
A;Residues: 19-27 <LED>
R;Meloun, B.; Moravek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A;Title: Complete amino acid sequence of human serum albumin.
A;Reference number: A91420; MUID:76187907; PMID:1225573
A;Accession: A91420
A;Molecule type: protein
A;Residues: 25-117,'EQ',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',395-
R;Roehr, U.; Spiteller, G.; Tripier, D.
Justus Liebigs Ann. Chem. 9, 881-884, 1988
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from ur
A;Reference number: S06422
A;Note: this paper is in German, with an English abstract
A;Accession: S06422
A;Molecule type: protein
A;Residues: 25-48 <ROE>
R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A;Title: Mass spectrometric identification of modifications to human serum albumin treat
A;Reference number: S36882; MUID:93384321; PMID:8373198
A;Accession: S36882
A;Molecule type: protein
A;Residues: 45-67;141-160;311-337;469-490;570-581 <FIN>
R;Kausler, E.; Spiteller, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A;Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol
A;Reference number: S17599; MUID:92126241; PMID:1772598
A;Accession: S17599
A;Molecule type: protein
A;Residues: 25-54;354-357;431-447 <KAU>
A;Note: 49-Leu was also found
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: A45800
A;Molecule type: protein
A;Residues: 166-173 <CAR>
R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre
A;Reference number: A03239; MUID:86242180; PMID:3087352
A;Accession: A03239
A;Molecule type: protein
A;Residues: 166-173,'L' <MOG>
R;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A;Title: Mutations in genetic variants of human serum albumin found in Italy.
A;Reference number: A38255; MUID:91062352; PMID:2247440
A;Accession: C38255
A;Molecule type: protein
A;Residues: 76-111 <GAL1>
A;Accession: B38255
A;Molecule type: protein
A;Residues: 82-105,'K',107-110 <GAL2>
A;Note: this variant is designated albumin Vibo Valentia
A;Accession: A38255
A;Molecule type: protein
A;Residues: 76-83,'K',85-106 <GAL3>
A;Note: this variant is designated albumin Torino
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A;Title: The structural characterization and bilirubin-binding properties of albumin Herb
A;Reference number: S33298; MUID:93292504; PMID:8513793
A;Accession: S33298
A;Molecule type: protein
A;Residues: 255-263,'E',265-281 <MIN1>
A;Note: this variant is designated albumin Herborn
R;Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,
Biochim. Biophys. Acta 1119, 232-238, 1992
A;Title: Two alloalbumins with identical electrophoretic mobility are produced by differe
A;Reference number: S21078; MUID:92190239; PMID:1347703
A;Accession: S21078
A;Molecule type: protein
A;Residues: 354-356,'K',358-378 <MIN2>
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported,
R;He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A;Title: Atomic structure and chemistry of human serum albumin.
A;Reference number: A46756; MUID:92334427; PMID:1630489
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
R;Brown, J.R.; Shockley, P.; Behrens, P.Q.
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,
A;Reference number: A94442
A;Contents: annotation; three-dimensional structure and disulfide bonds
R;Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A;Title: Disulfide bonds in human serum albumin.
A;Reference number: A90930
A;Contents: annotation; disulfide bonds
R;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding c
A;Reference number: A90299; MUID:78186630; PMID:656055
A;Contents: annotation; bilirubin-binding site
R;Peters, T.; Reed, R.G.
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjoholm, I., eds., 11-20,
A;Title: Serum albumin: conformation and active sites.
A;Reference number: A94408
A;Contents: annotation; binding sites
R;Harper, M.E.; Dugaiczyk, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes
A;Reference number: A90028; MUID:83279982; PMID:6192711
A;Contents: annotation; gene position
R;Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicyclic acid.
A;Reference number: A46755; MUID:76257808; PMID:955075
A;Contents: annotation
A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicyclic acid)
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphat
A;Reference number: A56294; MUID:92183881; PMID:1544460
A;Contents: annotation
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in pl
atase activity
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized i
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak b
C;Comment: A large number of variants of human serum albumin have been described.
C;Genetics:
A;Gene: GDB:ALB
A;Cross-references: GDB:118990; OMIM:103600
A;Map position: 4q11-4q13
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridoxa
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-609/Product: serum albumin #status experimental <MPT>
F;29-202/Domain: serum albumin repeat homology <SA1>

F;166-174/Product: kinetensin #status experimental <KIP>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4
F;214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 123; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.7e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
||||||||||||||||||||||||
Db 25 DAHKSEVAHRFKDLGEENFKALVL 48

RESULT 2
A47391
serum albumin precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A47391
R;Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A;Title: cDNA and protein sequence of polymorphic macaque albumins that differ in biliru
A;Reference number: A47391; MUID:93211971; PMID:8460152
A;Contents: B/B homozygote
A;Accession: A47391
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-600 <WAT>
A;Cross-references: UNIPROT:Q28522; GB:M90463; NID:g342294; PIDN:AAA36906.1; PID:g342295
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIP:128281)
C;Superfamily: serum albumin; serum albumin repeat homology
F;21-194/Domain: serum albumin repeat homology <SA1>
F;213-386/Domain: serum albumin repeat homology <SA2>
F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 89.4%; Score 110; DB 2; Length 600;
Best Local Similarity 87.5%; Pred. No. 3.6e-08;
Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
| |||||||||||||||:|| |||
Db 17 DTHKSEVAHRFKDLGEEHFKGLVL 40

RESULT 3
ABBOS
serum albumin precursor [validated] - bovine
N;Alternate names: 67K protein; preproalbumin
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94
R;Holowachuk, E.W.; Stoltenborg, J.K.; Reed, R.G.; Peters Jr., T.
submitted to the EMBL Data Library, August 1991
A;Description: Bovine serum albumin: cDNA sequence and expression.
A;Reference number: A38885
A;Accession: A38885
A;Molecule type: mRNA
A;Residues: 1-607 <HOL>
A;Cross-references: UNIPROT:P04277; EMBL:M73215
R;Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
Biochem. Biophys. Res. Commun. 173, 639-646, 1990
A;Title: Rapid confirmation and revision of the primary structure of bovine serum albumi
A;Reference number: A36401; MUID:91083649; PMID:2260975
A;Accession: A36401
A;Molecule type: protein
A;Residues: 25-41,'H',43-189,'E',191-213,'T',215-323,'D',325-393,'TS',396-607 <HIR>
R;MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
Eur. J. Biochem. 98, 477-485, 1979
A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A;Reference number: A91258; MUID:80024278; PMID:488109
A;Accession: A91258
A;Molecule type: protein
A;Residues: 1-32 <MAG>
R;Hsieh, J.C.; Lin, F.P.; Tam, M.F.
Anal. Biochem. 170, 1-8, 1988
A;Title: Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing ge
A;Reference number: A60808; MUID:88267456; PMID:3389500
A;Accession: B60808
A;Molecule type: protein
A;Residues: 25-41 <HSI>
R;Strawich, E.; Glimcher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albun
A;Reference number: S10780; MUID:90336641; PMID:2379503
A;Accession: S10780
A;Molecule type: protein
A;Residues: 25-41,'H',43-57,59-64 <STR>
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid protease
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: D45800
A;Molecule type: protein
A;Residues: 163-172 <CAR>
R;Carraway, R.E.; Mitra, S.P.; Cochrane, D.E.
J. Biol. Chem. 262, 5968-5973, 1987
A;Title: Structure of a biologically active neurotensin-related peptide obtained from pep
A;Reference number: A26693; MUID:87194805; PMID:2437111
A;Accession: A26693
A;Molecule type: protein
A;Residues: 165-172,'L' <CA2>
R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.
Biochem. J. 191, 867-868, 1980
A;Title: Sequence of residues 400-403 of bovine serum albumin.
A;Reference number: A90309; MUID:82023364; PMID:7283978
A;Accession: A90309
A;Molecule type: protein
A;Residues: 402-433 <REE>
R;Brown, J.R.
Fed. Proc. 34, 591, 1975
A;Title: Structure of bovine serum albumin.
A;Reference number: A91458
A;Accession: A91458
A;Molecule type: protein
A;Residues: 25-41,'H',43-117,'EQ',120-179,181-189,'E',191-194,'A',196-213,'T',215-288,'B
R;Brown, J.R.
submitted to the Atlas, April 1975
A;Reference number: A94551
A;Accession: A94551
A;Molecule type: protein
A;Residues: 190-195 <BR2>
R;Brown, J.R.
Fed. Proc. 33, 1389, 1974
A;Reference number: A91457
A;Contents: annotation; disulfide bonds
R;Werlen, R.C.; Offord, R.E.; Rose, K.
Biochem. J. 302, 907-911, 1994
A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC
A;Reference number: S55232; MUID:95031935; PMID:7945219
A;Accession: S55232
A;Status: preliminary
A;Molecule type: protein
A;Residues: 529-536;569-572 <WER>
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; copper binding; duplication; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-607/Product: serum albumin #status experimental <MPT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted

```
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4

  Query Match             87.8%;  Score 108;  DB 1;  Length 607;
  Best Local Similarity   83.3%;  Pred. No. 7.2e-08;
  Matches   20;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||:|||||||||| || |||
Db         25 DTHKSEIAHRFKDLGEEQFKGLVL 48


RESULT 4
ABRTS
serum albumin precursor - rat
N;Alternate names: preproalbumin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 09-Jul-2004
C;Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R;Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A;Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A;Reference number: A93872; MUID:81223722; PMID:7017712
A;Accession: A93872
A;Molecule type: mRNA
A;Residues: 1-608 <SAR>
A;Cross-references: UNIPROT:P02770; GB:V01222; GB:J00698; NID:g55627; PIDN:CAA24532.1; P
R;Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A;Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analys
A;Reference number: A92211; MUID:77249657; PMID:893447
A;Note: cleavages during protein maturation
A;Accession: A92211
A;Molecule type: protein
A;Residues: 1-38 <STR>
R;Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A;Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavag
A;Reference number: A91946; MUID:78109429; PMID:564345
A;Accession: A91946
A;Molecule type: protein
A;Residues: 25-222 <IS1>
R;Isemura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A;Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino a
A;Reference number: A91940; MUID:76260153; PMID:956149
A;Accession: A91940
A;Molecule type: protein
A;Residues: 223-288;572-608 <IS2>
A;Note: 262-Leu was also found
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.
A;Reference number: A90758; MUID:79001617; PMID:80265
A;Contents: annotation; copper binding
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: C45800
A;Status: preliminary
A;Molecule type: protein
A;Residues: 166-173 <CAR>
R;Heard, J.
Mol. Cell. Biol. 7, 2425-2434, 1987
A;Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved
A;Reference number: I57621; MUID:87286876; PMID:3475566
A;Accession: I57621
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: GB:M16825; NID:g202828; PIDN:AAA40712.1; PID:g554412
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-608/Product: serum albumin #status experimental <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status experimental
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,41

  Query Match             86.2%;  Score 106;  DB 1;  Length 608;
  Best Local Similarity   79.2%;  Pred. No. 1.4e-07;
  Matches   19;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :|||||:||||||||::|| |||
Db         25 EAHKSEIAHRFKDLGEQHFKGLVL 48


RESULT 5
ABSHS
serum albumin precursor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S06936
R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A;Reference number: S06936; MUID:90098888; PMID:2602160
A;Accession: S06936
A;Molecule type: mRNA
A;Residues: 1-607 <BRO>
A;Cross-references: UNIPROT:P14639; EMBL:X17055; NID:g1386; PIDN:CAA34903.1; PID:g1387
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, r
teroid hormones (weak bonds with these hormones promote their transfer across the membrar
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: serum albumin #status predicted <MAT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,41
F;263/Binding site: bilirubin (Lys) #status predicted

  Query Match             85.4%;  Score 105;  DB 1;  Length 607;
  Best Local Similarity   79.2%;  Pred. No. 2e-07;
  Matches   19;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||:|||| |||||||: |||
Db         25 DTHKSEIAHRFNDLGEENFQGLVL 48


RESULT 6
ABHOS
serum albumin precursor - horse
C;Species: Equus caballus (domestic horse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S34053
R;Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A;Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm r
A;Reference number: S34053; MUID:93345495; PMID:8344282
A;Accession: S34053
A;Molecule type: mRNA
A;Residues: 1-607 <HOA>
A;Cross-references: UNIPROT:P35747; GB:X74045; NID:g399671; PIDN:CAA52194.1; PID:g399672
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, r
teroid hormones (weak bonds with these hormones promote their transfer across the membrar
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
```

```
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: serum albumin #status predicted <MAT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F;263/Binding site: bilirubin (Lys) #status predicted

  Query Match             81.3%;  Score 100;  DB 1;  Length 607;
  Best Local Similarity   75.0%;  Pred. No. 1.1e-06;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | ||||:|||| |||||::|| |||
Db         25 DTHKSEIAHRFNDLGEKHFKGLVL 48


RESULT 7
ABPGS
serum albumin precursor - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S01382; A61006
R;Weinstock, J.; Baldwin, G.S.
Nucleic Acids Res. 16, 9045, 1988
A;Title: Nucleotide sequence of porcine liver albumin.
A;Reference number: S01382; MUID:89016582; PMID:3174440
A;Accession: S01382
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-605 <WEI>
A;Cross-references: UNIPROT:P08835; EMBL:X12422; NID:g1875; PIDN:CAA30970.1; PID:g833798
R;Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
J. Bone Miner. Res. 4, 235-241, 1989
A;Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral
A;Reference number: A61006; MUID:89269769; PMID:2728927
A;Accession: A61006
A;Molecule type: protein
A;Residues: 23-51,'X',53-54;'XXXGY',146,'E',148,'E',150-151,'XVN',155 <LIM>
A;Experimental source: dental enamel
A;Note: albumin and other serum proteins are also found in bone
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-16/Domain: signal sequence (fragment) #status predicted <SIG>
F;17-22/Domain: propeptide #status predicted <PRO>
F;23-605/Product: serum albumin #status predicted <MAT>
F;27-199/Domain: serum albumin repeat homology <SA1>
F;218-391/Domain: serum albumin repeat homology <SA2>
F;410-589/Domain: serum albumin repeat homology <SA3>
F;75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,4
F;261/Binding site: bilirubin (Lys) #status predicted

  Query Match             78.9%;  Score 97;  DB 1;  Length 605;
  Best Local Similarity   75.0%;  Pred. No. 2.9e-06;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              | :|||:|||||||||: || |||
Db         23 DTYKSEIAHRFKDLGEQYFKGLVL 46


RESULT 8
S57632
serum albumin precursor - cat
C;Species: Felis silvestris catus (domestic cat)
C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: JC4660; S57632
R;Hilger, C.; Grigioni, F.; Hentges, F.
Gene 169, 295-296, 1996
A;Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A;Reference number: JC4660; MUID:96194824; PMID:8647469
A;Accession: JC4660
A;Molecule type: mRNA
A;Residues: 1-608 <HI2>
A;Cross-references: UNIPROT:P49064; EMBL:X84842; NID:g886484; PIDN:CAA59279.1; PID:g88648
A;Experimental source: liver
C;Comment: This protein is the major protein component in plasma. It functions as a multi
ein has 35 conserved cysteine residues.
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: liver; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRP>
F;25-608/Product: serum albumin #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>

  Query Match             78.9%;  Score 97;  DB 2;  Length 608;
  Best Local Similarity   70.8%;  Pred. No. 2.9e-06;
  Matches   17;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :||:||:|||| |||||:|: |||
Db         25 EAHQSEIAHRFNDLGEEHFRGLVL 48


RESULT 9
JC5838
albumin - Mongolian jird
C;Species: Meriones unguiculatus (Mongolian jird)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5838
R;Yoshida, K.; Seto-Ohshima, A.; Sinohara, H.
DNA Res. 4, 351-354, 1997
A;Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the
A;Reference number: JC5838; MUID:98116663; PMID:9455485
A;Accession: JC5838
A;Molecule type: mRNA
A;Residues: 1-609 <YOS>
A;Cross-references: UNIPROT:O35090; DDBJ:AB006197; NID:g2317277; PIDN:BAA21765.1; PID:g2
A;Experimental source: liver
C;Superfamily: serum albumin; serum albumin repeat homology
F;222-395/Domain: serum albumin repeat homology <SA2>

  Query Match             78.9%;  Score 97;  DB 2;  Length 609;
  Best Local Similarity   78.3%;  Pred. No. 2.9e-06;
  Matches   18;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AHKSEVAHRFKDLGEENFKALVL 24
              |||||:|||:|||||: || |||
Db         27 AHKSEIAHRYKDLGEKYFKGLVL 49


RESULT 10
S29749
serum albumin - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C;Accession: S29749
R;Dixon, J.W.; Sarkar, B.
J. Biol. Chem. 249, 5872-5877, 1974
A;Title: Isolation, amino acid sequence and copper(II)-binding properties of peptide (1-
A;Reference number: S29749; MUID:75011422; PMID:4414013
A;Accession: S29749
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-24 <DIX>
C;Superfamily: serum albumin; serum albumin repeat homology

  Query Match             74.8%;  Score 92;  DB 2;  Length 24;
  Best Local Similarity   66.7%;  Pred. No. 5.7e-07;
  Matches   16;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              :|:|||:|||: |||||:|: |||
Db          1 EAYKSEIAHRYNDLGEEHFRGLVL 24


RESULT 11
B61511
serum albumin, milk-derived - Australian echidna (fragment)
C;Species: Tachyglossus aculeatus (Australian echidna)
C;Date: 09-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 09-Jul-2004
C;Accession: B61511
R;Teahan, C.G.; McKenzie, H.A.; Griffiths, M.
Comp. Biochem. Physiol. B 99, 99-118, 1991
A;Title: Some monotreme milk "whey" and blood proteins.
A;Reference number: A61511; MUID:92070088; PMID:1959333
A;Accession: B61511
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-30 <GRI>
A;Cross-references: UNIPROT:Q7M3A0
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: milk

  Query Match             72.4%;  Score 89;  DB 2;  Length 30;
  Best Local Similarity   66.7%;  Pred. No. 2e-06;
  Matches   16;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              || |||: ||:|:|||::|||| |
Db          1 DAQKSELGHRYKELGEDHFKALAL 24


RESULT 12
ABCHS
serum albumin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S15571; A05078; A13451
R;Cassady, A.I.; Salklld, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A;Reference number: S15571
A;Accession: S15571
A;Molecule type: mRNA
A;Residues: 1-615 <CAS>
A;Cross-references: UNIPROT:P19121; EMBL:X60688; NID:g63747; PIDN:CAA43098.1; PID:g63748
R;Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A;Title: The 5' noncoding and flanking regions of the avian very low density apolipoprot
A;Reference number: A05078; MUID:83161037; PMID:6187737
A;Accession: A05078
A;Molecule type: DNA
A;Residues: 1-28 <HAC>
A;Cross-references: GB:V00381; NID:g63038; PIDN:CAA23680.1; PID:g63039
R;Rosen, A.M.; Geller, D.M.
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A;Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.
A;Reference number: A13451; MUID:78019943; PMID:911327
A;Accession: A13451
A;Molecule type: protein
A;Residues: 19-23,'M',25-30 <ROS>
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), thyr
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-26/Domain: propeptide #status predicted <PRO>
F;27-613/Product: serum albumin #status predicted <MAT>
F;32-206/Domain: serum albumin repeat homology <SA1>
F;225-398/Domain: serum albumin repeat homology <SA2>
F;417-596/Domain: serum albumin repeat homology <SA3>
F;30/Binding site: copper (His) #status predicted
F;80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-397,

  Query Match             65.9%;  Score 81;  DB 1;  Length 615;
  Best Local Similarity   63.6%;  Pred. No. 0.00063;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKALVL 24
              ||||:|||: || || |||: :
Db         30 HKSEIAHRYNDLKEETFKAVAM 51


RESULT 13
T12251
sucrose synthase (EC 2.4.1.13) - common ice plant (fragment)
C;Species: Mesembryanthemum crystallinum (common ice plant)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C;Accession: T12251
R;Michalowski, C.B.; Bohnert, H.J.
submitted to the EMBL Data Library, March 1998
A;Description: An expressed sequence tag for sucrose synthase from M. crystallinum.
A;Reference number: Z17473
A;Accession: T12251
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-67 <MIC>
A;Cross-references: UNIPROT:O65179; EMBL:AF054446; NID:g3064040; PID:g3064041
C;Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             40.7%;  Score 50;  DB 2;  Length 67;
  Best Local Similarity   52.9%;  Pred. No. 2.1;
  Matches    9;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEE 17
              | | ||  |:|::|| |
Db         23 DTHYSEFEHKFQELGLE 39


RESULT 14
G85895
hypothetical protein yfgB [imported] - Escherichia coli (strain O157:H7, substrain EDL933
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85895
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85895
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-384 <STO>
A;Cross-references: UNIPROT:Q8XAA4; GB:AE005174; NID:g12516911; PIDN:AAG57627.1; GSPDB:GN
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: yfgB
C;Superfamily: conserved hypothetical protein HI0365

  Query Match             40.7%;  Score 50;  DB 2;  Length 384;
  Best Local Similarity   42.9%;  Pred. No. 13;
  Matches    9;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKA 21
              | :: ::   ||||||: |:|
Db         22 DLNRQQMREFFKDLGEKPFRA 42


RESULT 15
C91051
hypothetical protein ECs3379 [imported] - Escherichia coli (strain O157:H7, substrain RIM
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: C91051
```

```
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: C91051
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-384 <HAY>
A;Cross-references: UNIPROT:Q8XAA4; GB:BA000007; PIDN:BAB36802.1; PID:g13362849; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs3379
C;Superfamily: conserved hypothetical protein HI0365

  Query Match             40.7%;  Score 50;  DB 2;  Length 384;
  Best Local Similarity   42.9%;  Pred. No. 13;
  Matches    9;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKA 21
              | :: ::   ||||||: |:|
Db         22 DLNRQQMREFFKDLGEKPFRA 42


RESULT 16
D65028
hypothetical 43.1 kD protein in ndk-gcpE intergenic region - Escherichia coli (strain K-
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: D65028
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cc
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: D65028
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-384 <BLAT>
A;Cross-references: UNIPROT:P36979; GB:AE000338; GB:U00096; NID:g1788862; PIDN:AAC75570.
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yfgB
C;Superfamily: conserved hypothetical protein HI0365

  Query Match             40.7%;  Score 50;  DB 2;  Length 384;
  Best Local Similarity   42.9%;  Pred. No. 13;
  Matches    9;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKA 21
              | :: ::   ||||||: |:|
Db         22 DLNRQQMREFFKDLGEKPFRA 42


RESULT 17
A53361
thyroxine-binding protein, plasma - turtle (Trachemys scripta) (fragments)
N;Alternate names: TBP protein
C;Species: Trachemys scripta (turtle)
C;Date: 19-Mar-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jul-2004
C;Accession: A53361; S41664
R;Licht, P.
submitted to the Protein Sequence Database, December 1993
A;Reference number: A53361
A;Accession: A53361
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-30;31-45;46-66;67-74;75-84;85-94;95-109;110-122;123-138;139-153 <LIC>
A;Cross-references: UNIPROT:Q7LZ08
R;Licht, P.; Moore, M.F.
Arch. Biochem. Biophys. 309, 47-51, 1994
A;Title: Structure of a reptilian plasma thyroxine binding protein indicates homology tc
A;Reference number: S41664; MUID:94161555; PMID:8117112
A;Accession: S41664
A;Molecule type: protein
A;Residues: 1-30;31-45;46-66;67-74;75-84;85-94;95-109;110-122;123-138;139-153 <LIW>
C;Keywords: plasma

  Query Match             39.8%;  Score 49;  DB 2;  Length 153;
  Best Local Similarity   33.3%;  Pred. No. 7;
  Matches    8;  Conservative    8;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKALVL 24
              |  : :|   | :||::||::| :
Db          6 DYVRDKVCQEFNNLGKDNFRSLAI 29


RESULT 18
S56295
SAP155 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein R013; protein YFR040w
C;Species: Saccharomyces cerevisiae
C;Date: 02-Sep-1995 #sequence_revision 19-Oct-1995 #text_change 20-Jun-2000
C;Accession: S56295; S62251; S63787
R;Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanun
submitted to the EMBL Data Library, May 1995
A;Description: Analysis of the nucleotide sequence of chromosome VI from Saccaromyces ce
A;Reference number: S56186
A;Accession: S56295
A;Molecule type: DNA
A;Residues: 1-905 <MUR>
A;Cross-references: EMBL:D50617; NID:g836685; PID:g836795; MIPS:YFR040w
R;Murakami, Y.
submitted to the EMBL Data Library, December 1994
A;Reference number: S62230
A;Accession: S62251
A;Molecule type: DNA
A;Residues: 1-905 <MUW>
A;Cross-references: EMBL:D44597; NID:g871938; PID:g871939
R;Eki, T.; Naitou, M.; Hagiwara, H.; Ozawa, M.; Sasanuma, S.I.; Sasanuma, M.; Tsuchiya,
Yeast 12, 149-167, 1996
A;Title: Analysis of a 36.2 kb DNA sequence including the right telomere of chromosome V
A;Reference number: S63787; MUID:96287652; PMID:8686379
A;Accession: S63787
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-905 <EKI>
A;Cross-references: EMBL:D44597; NID:g871938; PID:g871939
C;Genetics:
A;Gene: SGD:SAP155
A;Cross-references: SGD:S0001936; MIPS:YFR040w
A;Map position: 6R

  Query Match             39.8%;  Score 49;  DB 2;  Length 905;
  Best Local Similarity   40.9%;  Pred. No. 43;
  Matches    9;  Conservative    6;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKALVL 24
              |::::  :|| || | || : |
Db        442 HENQLHEKFKPLGFERFKVVEL 463


RESULT 19
T20006
hypothetical protein C47E12.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T20006
R;Coles, L.
submitted to the EMBL Data Library, January 1996
A;Reference number: Z19210
A;Accession: T20006
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-184 <WIL>
A;Cross-references: UNIPROT:Q18676; EMBL:Z68882; PIDN:CAA93103.1; GSPDB:GN00022; CESP:C4
```

```
A;Experimental source: clone C47E12
C;Genetics:
A;Gene: CESP:C47E12.9
A;Map position: 4
A;Introns: 25/3; 51/3; 115/3
C;Superfamily: Caenorhabditis elegans hypothetical protein C47E12.9

  Query Match             39.0%;  Score 48;  DB 2;  Length 184;
  Best Local Similarity   60.0%;  Pred. No. 12;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEE 17
              : |||| :|:|| ||
Db        149 YSSEVAQKFEDLSEE 163


RESULT 20
A97575
hypothetical protein AGR_C_3280 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: A97575
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A97575
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <KUR>
A;Cross-references: UNIPROT:Q8UEH3; GB:AE007869; PIDN:AAK87554.1; PID:g15156890; GSPDB:G
C;Genetics:
A;Gene: AGR_C_3280
A;Map position: circular chromosome
C;Superfamily: Neisseria meningitidis probable integral membrane protein NMA2020

  Query Match             39.0%;  Score 48;  DB 2;  Length 222;
  Best Local Similarity   50.0%;  Pred. No. 14;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 KSEVAHRFKDLGEE 17
              | |:|||| ::|::
Db        151 KGELAHRFNEMGQQ 164


RESULT 21
AI2795
conserved hypothetical protein Atu1784 [imported] - Agrobacterium tumefaciens (strain C5
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AI2795
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AI2795
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <KUR>
A;Cross-references: UNIPROT:Q8UEH3; GB:AE008688; PIDN:AAL42783.1; PID:g17740227; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu1784
A;Map position: circular chromosome
C;Superfamily: Neisseria meningitidis probable integral membrane protein NMA2020

  Query Match             39.0%;  Score 48;  DB 2;  Length 222;
  Best Local Similarity   50.0%;  Pred. No. 14;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 KSEVAHRFKDLGEE 17
              | |:|||| ::|::
Db        151 KGELAHRFNEMGQQ 164


RESULT 22
VYHUD
vitamin D-binding protein precursor [validated] - human
N;Alternate names: DBP; Gc-globulin; group-specific component
C;Species: Homo sapiens (man)
C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 09-Jul-2004
C;Accession: A94076; A46759; A29096; A92765; S39787; A24066; A90427; A03237
R;Yang, F.; Brune, J.L.; Naylor, S.L.; Cupples, R.L.; Naberhaus, K.H.; Bowman, B.H.
Proc. Natl. Acad. Sci. U.S.A. 82, 7994-7998, 1985
A;Title: Human group-specific component (Ge) is a member of the albumin family.
A;Reference number: A94076; MUID:86068030; PMID:2415977
A;Accession: A94076
A;Molecule type: mRNA
A;Residues: 1-474 <YAN1>
A;Cross-references: UNIPROT:P02774; GB:X03178; GB:M11321; NID:g31675; PIDN:CAA26938.1; PI
A;Experimental source: allele Gc2
R;Witke, W.F.; Gibbs, P.E.M.; Zielinski, R.; Yang, F.; Bowman, B.H.; Dugaiczyk, A.
Genomics 16, 751-754, 1993
A;Title: Complete structure of the human Gc gene: differences and similarities between me
A;Reference number: A46759; MUID:93315171; PMID:8325650
A;Accession: A46759
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-431,'E',433-435,'T',437-474 <WIT>
A;Cross-references: GB:L10641; NID:g340281; PIDN:AAA61704.1; PID:g639896
A;Experimental source: allele Gc1
R;Yang, F.; Naberhaus, K.H.; Adrian, G.S.; Gardella, J.M.; Brissenden, J.E.; Bowman, B.H.
Gene 54, 285-290, 1987
A;Title: The vitamin D-binding protein gene contains conserved nucleotide sequences that
A;Reference number: A29096; MUID:88005794; PMID:2958390
A;Accession: A29096
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-19 <YAN2>
A;Cross-references: GB:M17156; NID:g181489; PIDN:AAA19662.1; PID:g463096
R;Cooke, N.E.; David, E.V.
J. Clin. Invest. 76, 2420-2424, 1985
A;Title: Serum vitamin D-binding protein is a third member of the albumin and alpha feto
A;Reference number: A92765; MUID:86086396; PMID:2416779
A;Accession: A92765
A;Molecule type: mRNA
A;Residues: 1-167,'E',169-326,'R',328-431,'E',433-435,'T',437-474 <COO>
A;Cross-references: GB:M12654; NID:g181481; PIDN:AAA52173.1; PID:g181482
A;Experimental source: allele Gc1
R;Braun, A.; Kofler, A.; Morawietz, S.; Cleve, H.
Biochim. Biophys. Acta 1216, 385-394, 1993
A;Title: Sequence and organization of the human vitamin D-binding protein gene.
A;Reference number: S39787; MUID:94092730; PMID:7505619
A;Accession: S39787
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-431,'E',433-435,'T',437-474 <BRA>
A;Cross-references: GB:S67480; NID:g455967; PIDN:AAB29423.1; PID:g455970
R;Schoentgen, F.; Metz-Boutigue, M.H.; Jolles, J.; Constans, J.; Jolles, P.
Biochim. Biophys. Acta 871, 189-198, 1986
A;Title: Complete amino acid sequence of human vitamin D-binding protein (group-specific
A;Reference number: A24066; MUID:86216223; PMID:2423133
A;Accession: A24066
A;Molecule type: protein
A;Residues: 17-474 <SCH>
R;Svasti, J.; Kurosky, A.; Bennett, A.; Bowman, B.H.
Biochemistry 18, 1611-1617, 1979
A;Title: Molecular basis for the three major forms of human serum vitamin D binding prote
A;Reference number: A90427; MUID:79145448; PMID:218624
A;Accession: A90427
A;Molecule type: protein
```

A;Residues: 17,'Q',19-21,'N',23-36,'XXX',40-41;472-474 <SVA>
C;Comment: DBP is a multifunctional protein found in plasma, ascitic fluid, cerebrospina
nts polymerization of actin by binding its monomers. DBP associates with membrane-bound
C;Comment: Over 80 variants of human DBP have been identified. The three most common all
C;Genetics:
A;Gene: GDB:GC
A;Cross-references: GDB:119263; OMIM:139200
A;Map position: 4q12-4q13
A;Introns: 20/1; 43/2; 87/3; 158/2; 202/3; 234/2; 277/3; 345/2; 388/3; 421/2; 465/3
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: actin binding; duplication; globulin; glycoprotein; liver; plasma; polymorph
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-474/Product: vitamin D-binding protein (allele Gc2) #status experimental <MAL2>
F;17-431,'E',433-435,'T',437-474/Product: vitamin D-binding protein (allele Gc1) #status
F;26-199/Domain: serum albumin repeat homology <SA1>
F;217-385/Domain: serum albumin repeat homology <SA2>
F;404-474/Domain: serum albumin repeat homology #status atypical <SA3>
F;29-75,74-83,96-112,111-122,145-190,189-198,220-266,265-273,286-300,299-311,335-376,375
F;288/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match 39.0%; Score 48; DB 1; Length 474;
  Best Local Similarity 41.7%; Pred. No. 31;
  Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENFKALVL 24
Db 22 DYEKNKVCKEFSHLGKEDFTSLSL 45

RESULT 23
T06675
hypothetical protein T17F15.50 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06675
R;Quetier, F.; Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigu
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15793
A;Accession: T06675
A;Molecule type: DNA
A;Residues: 1-629 <QUE>
A;Cross-references: UNIPROT:Q9SU71; EMBL:AL049658; GSPDB:GN00061; ATSP:T17F15.50
A;Experimental source: cultivar Columbia; BAC clone T17F15
C;Genetics:
A;Gene: ATSP:T17F15.50
A;Map position: 3
A;Introns: 106/3; 357/3
C;Superfamily: Arabidopsis thaliana disease resistance protein EDS1

  Query Match 38.6%; Score 47.5; DB 2; Length 629;
  Best Local Similarity 56.5%; Pred. No. 49;
  Matches 13; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

Qy 2 AHKSEVAHRFKDLGEEN-FKALV 23
Db 406 AHKNGYYDSFKDSNEENDFKANV 428

RESULT 24
AB0230
probable exported protein YPO1886 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AB0230
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AB0230
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-189 <KUR>
A;Cross-references: UNIPROT:Q8ZF35; GB:AL590842; PIDN:CAC90702.1; PID:g15979905; GSPDB:GN
C;Genetics:
A;Gene: YPO1886

  Query Match 38.2%; Score 47; DB 2; Length 189;
  Best Local Similarity 50.0%; Pred. No. 17;
  Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEEN 18
Db 51 DSEKSEIFKQFKNLLDEN 68

RESULT 25
E83095
conserved hypothetical protein PA4420 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83095
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83095
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-313 <STO>
A;Cross-references: UNIPROT:Q9HVZ5; GB:AE004856; GB:AE004091; NID:g9950633; PIDN:AAG0780
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA4420
C;Superfamily: Escherichia coli yabC protein

  Query Match 38.2%; Score 47; DB 2; Length 313;
  Best Local Similarity 47.4%; Pred. No. 28;
  Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DAHKSEVAHRFKDLGEENF 19
Db 144 NATEEEIARVFKEYGEERF 162

RESULT 26
F96577
hypothetical protein F22G10.3 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 16-Aug-2004
C;Accession: F96577
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96577
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-719 <STO>
A;Cross-references: UNIPROT:Q9C8M9; GB:AE005173; NID:g10645349; PIDN:AAG21469.1; GSPDB:G
C;Genetics:
A;Gene: F22G10.3
A;Map position: 1
C;Superfamily: protein kinase homology

  Query Match 38.2%; Score 47; DB 2; Length 719;

```
  Best Local Similarity   52.4%;  Pred. No. 67;
  Matches   11;  Conservative    1;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKALV 23
              ||:   | |  | ||  ||||
Db        494 HKNGSLHDFLHLSEEESKALV 514


RESULT 27
F89130
protein F52E1.10 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: F89130
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: F89130
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-451 <STO>
A;Cross-references: UNIPROT:Q20666; GB:chr_V; PIDN:AAB37043.1; PID:g1086810; GSPDB:GN000
C;Genetics:
A;Gene: F52E1.10
A;Map position: 5

  Query Match             37.8%;  Score 46.5;  DB 2;  Length 451;
  Best Local Similarity   48.1%;  Pred. No. 49;
  Matches   13;  Conservative    1;  Mismatches    8;  Indels    5;  Gaps    1;

Qy          2 AHKSEV-----AHRFKDLGEENFKALV 23
              ||| ||     |||  |  :|  | ||
Db        349 AHKCEVFWNENAHRLNDNRQELLKLLV 375


RESULT 28
T44944
hypothetical protein 5 [imported] - Natronobacterium pharaonis
C;Species: Natronobacterium pharaonis
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T44944
R;Mattar, S.; Engelhard, M.
Eur. J. Biochem. 250, 332-341, 1997
A;Title: Cytochrome ba3 from Natronobacterium pharaonis: An archaeal four-subunit cytoch
A;Reference number: Z22876; MUID:98088958; PMID:9428682
A;Accession: T44944
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-180 <MAT>
A;Cross-references: UNIPROT:O07291; EMBL:Y10500; PIDN:CAA71527.1
A;Experimental source: strain SP1/28
C;Genetics:
A;Note: orf5
C;Superfamily: conserved hypothetical protein AF1745

  Query Match             37.4%;  Score 46;  DB 2;  Length 180;
  Best Local Similarity   56.2%;  Pred. No. 23;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGE 16
              | | || | ||: :||
Db        100 DYHASEYASRFESVGE 115


RESULT 29
B72219
ribulose-phosphate 3-epimerase - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: B72219
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequ
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: B72219
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-220 <ARN>
A;Cross-references: UNIPROT:Q9X243; GB:AE001811; GB:AE000512; NID:g4982291; PIDN:AAD36784
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1718
C;Superfamily: yeast ribulose-5-phosphate-epimerase

  Query Match             37.4%;  Score 46;  DB 2;  Length 220;
  Best Local Similarity   45.0%;  Pred. No. 28;
  Matches    9;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKAL 22
              |   : || ||||   | |:
Db         94 HLHRLVHRVKDLGARAFVAI 113


RESULT 30
G95324
probable TraD conjugal transfer protein [imported] - Sinorhizobium meliloti (strain 1021)
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: G95324
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowse
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.,
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: G95324
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-79 <KUR>
A;Cross-references: UNIPROT:Q92ZI2; GB:AE006469; PIDN:AAK65161.1; PID:g14523604; GSPDB:GN
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: traD
A;Genome: plasmid

  Query Match             36.6%;  Score 45;  DB 2;  Length 79;
  Best Local Similarity   60.0%;  Pred. No. 14;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         10 RFKDLGEENFKALVL 24
              | :|:||| |||  |
Db         51 RLRDIGEEAFKAPAL 65


RESULT 31
T36873
hypothetical protein SCI51.10c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36873
R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21617
```

```
A;Accession: T36873
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-264 <MUR>
A;Cross-references: UNIPROT:Q9S230; EMBL:AL109848; PIDN:CAB52834.1; GSPDB:GN00070; SCOED
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCI51.10c
C;Superfamily: Streptomyces coelicolor hypothetical protein SCI51.10c

  Query Match             36.6%;  Score 45;  DB 2;  Length 264;
  Best Local Similarity   52.9%;  Pred. No. 47;
  Matches    9;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          7 VAHRFKDLGEENFKALV 23
              |  | |:||| :|: ||
Db        180 VVARLKELGEASFRVLV 196


RESULT 32
S56055
hypothetical protein YMR241w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YM9408.03
C;Species: Saccharomyces cerevisiae
C;Date: 27-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: S56055
R;Gentles, S.; Bowman, S.
submitted to the EMBL Data Library, March 1995
A;Reference number: S56053
A;Accession: S56055
A;Molecule type: DNA
A;Residues: 1-314 <GEN>
A;Cross-references: UNIPROT:Q04013; EMBL:Z48756; NID:g736304; PID:g736307; GSPDB:GN00013
C;Genetics:
A;Gene: SGD:YHM2; MIPS:YMR241w
A;Cross-references: SGD:S0004854
A;Map position: 13R
C;Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

  Query Match             36.6%;  Score 45;  DB 2;  Length 314;
  Best Local Similarity   45.0%;  Pred. No. 56;
  Matches    9;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 SEVAHRFKDLGEENFKALVL 24
              :|  :||| ||  || : :|
Db         96 AEAEYRFKSLGLNNFASGIL 115


RESULT 33
G64667
NA+/H+ antiporter - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: G64667
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: G64667
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-383 <TOM>
A;Cross-references: UNIPROT:O25795; GB:AE000624; GB:AE000511; NID:g2314340; PIDN:AAD0822
C;Superfamily: Aquifex aeolicus Na+/H+-exchanging protein napA1

  Query Match             36.6%;  Score 45;  DB 2;  Length 383;
  Best Local Similarity   53.8%;  Pred. No. 69;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLG 15
              ||||: |:  |:|
Db        254 HKSELIHKLNDVG 266


RESULT 34
H71848
probable na+/h+ antiporter - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: H71848
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: H71848
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-383 <ARN>
A;Cross-references: UNIPROT:Q9ZK33; GB:AE001538; GB:AE001439; NID:g4155697; PIDN:AAD0668
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp1109
C;Superfamily: Aquifex aeolicus Na+/H+-exchanging protein napA1

  Query Match             36.6%;  Score 45;  DB 2;  Length 383;
  Best Local Similarity   53.8%;  Pred. No. 69;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLG 15
              ||||: |:  |:|
Db        254 HKSELIHKLNDVG 266


RESULT 35
AD0822
conserved hypothetical protein STY2770 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AD0822
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AD0822
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-388 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD02728.1; PID:g16503740; GSPDB:GN00176
C;Genetics:
A;Gene: STY2770
C;Superfamily: conserved hypothetical protein HI0365

  Query Match             36.6%;  Score 45;  DB 2;  Length 388;
  Best Local Similarity   38.1%;  Pred. No. 69;
  Matches    8;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEENFKA 21
              | :: ::   ||:|||: |:|
Db         26 DLNRQQMREFFKNLGEKPFRA 46


RESULT 36
S41952
UTP-glucose glucosyltransferase - cassava (fragment)
C;Species: Manihot esculenta (cassava)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S41952
```

```
R;Hughes, J.; Hughes, M.A.
submitted to the EMBL Data Library, January 1994
A;Description: Multiple secondery plant product UDP-glucose glucosyltransferase genes ex
A;Reference number: S41950
A;Accession: S41952
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-394 <HUG>
A;Cross-references: UNIPROT:Q40288; EMBL:X77463; NID:g453250; PID:g453251
C;Superfamily: flavonol O3-glucosyltransferase

  Query Match             36.6%;  Score 45;  DB 2;  Length 394;
  Best Local Similarity   38.1%;  Pred. No. 71;
  Matches    8;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 HKSEVAHRFKDLGEENFKALV 23
              | ||:  | |:: ::: |||:
Db        353 HHSEIRKRVKEMSDKSRKALM 373


RESULT 37
A45565
cysteine proteinase (EC 3.4.22.-) - Theileria annulata
C;Species: Theileria annulata
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A45565
R;Baylis, H.A.; Megson, A.; Mottram, J.C.; Hall, R.
Mol. Biochem. Parasitol. 54, 105-107, 1992
A;Title: Characterisation of a gene for a cysteine protease from Theileria annulata.
A;Reference number: A45565; MUID:92389980; PMID:1518523
A;Accession: A45565
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-441 <BAY>
A;Cross-references: UNIPROT:P25781; GB:M86659; NID:g161886; PID:g161887
A;Experimental source: Hissar
A;Note: sequence extracted from NCBI backbone (NCBIN:112709, NCBIP:112710)
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase
F;252,381,403/Active site: Cys, His, Asn #status predicted

  Query Match             36.6%;  Score 45;  DB 2;  Length 441;
  Best Local Similarity   57.1%;  Pred. No. 79;
  Matches    8;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          9 HRFKDLGEENFKAL 22
              ::| || :| ||||
Db        165 NKFSDLSDEEFKAL 178


RESULT 38
A97208
carbon monoxide dehydrogenase, catalytic chain (cooS) [imported] - Clostridium acetobuty
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 15-Mar-2004
C;Accession: A97208
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: A97208
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-639 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80452.1; PID:g15025519; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2498
C;Superfamily: carbon-monoxide dehydrogenase, beta subunit; hybrid cluster [4Fe-2S-3O] h

  Query Match             36.6%;  Score 45;  DB 2;  Length 639;
  Best Local Similarity   52.6%;  Pred. No. 1.2e+02;
  Matches   10;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 AHKSEVAHRFKDLGEENFK 20
              |   | | |  :|| ||||
Db        383 AQADETAQRLIELGIENFK 401


RESULT 39
C82860
DNA gyrase subunit B XF0005 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: C82860
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82860
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-814 <SIM>
A;Cross-references: UNIPROT:Q9PHD9; GB:AE003855; GB:AE003849; NID:g9104760; PIDN:AAF82818
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Za
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0005
C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

  Query Match             36.6%;  Score 45;  DB 2;  Length 814;
  Best Local Similarity   55.6%;  Pred. No. 1.5e+02;
  Matches   10;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DAHKSEVAHRFKDLGEEN 18
              || |     ||| ||| |
Db        739 DAKKGRQIQRFKGLGEMN 756


RESULT 40
T18552
saframycin Mx1 synthetase A - Myxococcus xanthus
C;Species: Myxococcus xanthus
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18552
R;Pospiech, A.; Bietenhader, J.; Schupp, T.
Microbiology 142, 741-746, 1996
A;Title: Two multifunctional peptide synthetases and an O-methyltransferase are involved
A;Reference number: Z18967; MUID:97090395; PMID:8936303
A;Accession: T18552
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2605 <POS>
A;Cross-references: UNIPROT:Q50858; EMBL:U24657; NID:g1171127; PID:g1171129; PIDN:AAC441
C;Genetics:
A;Gene: safA
C;Keywords: carrier protein
F;535-978/Domain: acetate-CoA ligase homology <ACL1>
F;997-1065/Domain: acyl carrier protein homology <ACP1>
F;1643-2091/Domain: acetate-CoA ligase homology <ACL2>
F;2110-2178/Domain: acyl carrier protein homology <ACP2>
```

Query Match 36.6%; Score 45; DB 2; Length 2605;
Best Local Similarity 46.7%; Pred. No. 4.9e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 DAHKSEVAHRFKDLG 15
              ||| :::||  : ||
Db       1629 DAHANQLAHHLRQLG 1643
```

Search completed: August 19, 2005, 10:58:46
Job time : 18 secs

THIS PAGE BLANK (USPTO)